Systems analysis of protein signatures predicting cetuximab responses in KRAS, NRAS, BRAF and PIK3CA wild-type patient-derived xenograft models of metastatic colorectal cancer

IJC International Journal of Cancer

Systems analysis of protein signatures predicting Cetuximab responses in KRAS, NRAS, BRAF and PIK3CA wild-type patient-derived xenografts models of metastatic colorectal cancer

Journal:	*International Journal of Cancer*
Manuscript ID	IJC-20-0833.R1
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Lindner, Andreas; Royal College of Surgeons in Ireland, Physiology and Medical Physics Carberry, Steven; Royal College of Surgeons in Ireland, Physiology and Medical Physics Monsefi, Naser; Royal College of Surgeons in Ireland, Physiology and Medical Physics Barat, Ana; Royal College of Surgeons in Ireland, Department of Physiology and Medical Physics Salvucci, Manuela; Royal College of Surgeons in Ireland, Physiology and Medical Physics O'Byrne, Robert; Royal College of Surgeons in Ireland, Physiology and Medical Physics Zanella, Eugenia; University of Torino Medical School, IRCC, Institute for Cancer Research and Treatment, Division of Molecular Oncology Cremona, Mattia; Royal College of Surgeons in Ireland, Department of Medical Oncology Hennessy, Bryan; Royal College of Surgeons in Ireland, Department of Medical Oncology, Molecular Medicine Laboratories Bertotti, Andrea; University of Torino Medical School, Laboratories of Molecular Pharmacology and Department of Oncoloty Trusolino, Livio; University of Torino Medical School, IRCC, Institute for Cancer Research and Treatment Prehn, Jochen; Royal College of Surgeons in Ireland, Department of Physiology and Medical Physics
Key Words:	metastatic, colorectal cancer, patient derived xenograpft, cetuximab

Systems analysis of protein signatures predicting Cetuximab responses in *KRAS*, *NRAS*, *BRAF* and *PIK3CA* wild-type patient-derived xenografts models of metastatic colorectal cancer

Andreas U. Lindner[1], Steven Carberry[1], Naser Monsefi[1], Ana Barat[1], Manuela Salvucci[1], Robert O'Byrne[1], Eugenia R. Zanella[2], Mattia Cremona[3], Bryan T. Hennessy[3], Andrea Bertotti[2,4], Livio Trusolino[2,4] and Jochen H.M. Prehn[1]

[1]Department of Physiology and Medical Physics and Centre Systems Medicine, Royal College of Surgeons in Ireland, Dublin, Ireland; [2]Translational Cancer Medicine, Surgical Oncology, and Clinical Trials Coordination, Candiolo Cancer Institute Fondazione del Piemonte per l'Oncologia IRCCS, Turin, Italy; [3]Department of Medical Oncology, Beaumont Hospital, Royal College of Surgeons in Ireland, Dublin, Ireland; [4]Department of Oncology, University of Turin Medical School, Turin, Italy.

Corresponding Author: Prof. Jochen H.M. Prehn, Centre for Systems Medicine, and Department of Physiology and Medical Physics, Royal College of Surgeons in Ireland, 123 St. Stephen's Green, Dublin 2, Ireland. Tel.: +353 (1) 402 2255. Fax: + 353 (1) 402 2447. eMail: prehn@rcsi.ie

Running title: Systems analysis of cetuximab responses

Novelty and Impact: A large fraction of patients with metastatic colorectal cancer do not respond to anti-EGFR therapy despite KRAS wild type tumours. Statistical analysis of RPPA data of colorectal cancer KRAS, BRAF, NRAS and PI3KCA wild type PDX models revealed a 14 - 20 (phospho)protein signature that was predicting responses to cetuximab. Our findings furthermore emphasise GSK-3β to be potentially targetable for a co-treatment with cetuximab.

Keywords: anti-EGFR, metastatic colorectal cancer, molecular subtyping, reverse-phase protein array, deterministic modelling, apoptosis, proliferation

Abbreviations: 5-FU, fluorouracil; ANOVA, analysis of variance; CMS, consensus molecular subtypes; CRC, colorectal cancer; CRIS, CRC intrinsic subtype; EGF, epidermal growth factor; EGFR, EGF receptor; LASSO, least absolute shrinkage and selection operator; NMF, non-negative matrix factorization; P, p-value; PAM, Prediction Analysis for Microarrays; PDX, patient-derived mouse xenograft; RPPA, reverse phase protein array; SC, substrate cleavage.

Abstract

Antibodies targeting the human epidermal growth factor receptor (*EGFR*) are used for the treatment of *RAS* wild-type metastatic colorectal cancer. A significant proportion of patients remains unresponsive to this therapy. Here, we performed a reverse phase protein array-based (phospho)protein analysis of 63 ~~'quadruple-negative' (~~*KRAS*, *NRAS*, *BRAF* and *PIK3CA* wild-type~~)~~ metastatic CRC tumours. Responses of tumours to anti-EGFR therapy with cetuximab were recorded in patient-derived xenograft (PDX) models. Unsupervised hierarchical clustering of pre-treatment tumour tissue identified three clusters, of which cluster C3 was exclusively composed of responders. Clusters C1 and C2 showed mixed responses. None of the three protein clusters showed a significant correlation with transcriptome-based subtypes. Analysis of protein signatures across all PDXs identified 14 markers that discriminated cetuximab-sensitive and -resistant tumours: PDK1 (S241), Caspase-8, Shc (Y317), Stat3 (Y705), p27, GSK-3β (S9), HER3, PKC-α (S657), EGFR (Y1068), Akt (S473), S6 Ribosomal Protein (S240/244), HER3 (Y1289), NF-κB-p65 (S536) and Gab-1 (Y627). Least absolute shrinkage and selection operator and binominal logistic regression analysis delivered refined protein signatures for predicting response to cetuximab. (Phospo-)protein analysis of matched pre- and post-treated models furthermore showed significant reduction of Gab-1 (Y627) and GSK-3β (S9) exclusively in responding models, suggesting novel targets for treatment.

Background

Colorectal cancer (CRC) is the third and second most commonly diagnosed cancer in males and females, and the second most common cause of cancer-related deaths in the developed world. In the advanced setting, CRC is routinely treated with fluorouracil (5-FU)-based chemotherapy. 30% of CRC patients present in the metastatic setting[1] where response rates to palliative 5-FU/oxaliplatin- or 5-FU/irinotecan-based chemotherapy range between 40-50%. Median overall survival remains poor at around 16-19 months[2]. Identifying the importance of epidermal growth factor (*EGF*) signalling for the survival of CRC cells resulted in the development of targeted therapies that neutralize the oncogenic activity of EGF receptors (*EGFR*). Anti-EGFR therapies have significantly improved survival in metastatic CRC patients[3]. Guidelines recommend to test for *KRAS*, *NRAS* and *BRAF* mutations as well as microsatellite instability status in CRC patients being considered for anti-EGFR therapy[4, 5] on the bases of the ineffectiveness of anti-EGFR therapy ~~is not effective~~ in patients with activating *KRAS*, *BRAF*, and *NRAS* mutations[6], and favourable responses to immune check point inhibitors in microsatellite instability-high patients[4]. While *PI3KCA* mutational analysis is not recommended yet[4], *PIK3CA* exon 20 mutations were linked with a worse outcome compared with wild-type status in patients with metastatic colorectal cancer [7]. Nevertheless, between 50–60% of patients will not benefit from anti-EGFR treatment even when these are *KRAS*, *BRAF*,

NRAS and *PI3KCA* wild type ~~(*quadruple negative*) have a 'quadruple negative' status~~[7].

Mutations and copy number alterations in genes encoding for other survival signaling proteins have been shown to contribute to anti-EGFR resistance. For example, *HER2*-amplification, *IGF2* overexpression or increased MET activity resulted in reduced responses to anti-EGFR therapy, as demonstrated in patient-derived xenograft (PDX) models of metastatic CRC and in patients[8, 9]. Analysis of the genomic and transcriptomic landscape of anti-EGFR resistance in PDX models and patients furthermore identified mutations in *EGFR*, *FGFR1*, *PDGFRA*, and *MAP2K1* or loss of *NF1* to contribute to anti-EGF resistance[9, 10].

While identification of patient-specific genome alterations provides a personalised diagnosis that provides insights into anti-EGFR therapy responses and may open opportunities for personalised therapies, interpretation of often multiple genomic alterations found in most patients is not always straightforward. Other efforts to identify responders and non-responders to anti-EGFR therapy have therefore focussed on the power of unsupervised molecular subtyping of tumours. An international meta-analysis and bioinformatics effort led to the identification of four distinct subtypes in CRC, termed 'Consensus Molecular Subtypes' (CMS1-CMS4)[11]. A recent study demonstrated that CMS2 patients benefitted more from anti-EGFR therapy than patients treated with anti-angiogenic therapy, while the opposite was the case in CMS1 patients[12]. However predictions of anti-EGFR therapy responses in CMS3 and CMS4 patients were not possible, and significant variability in overall and progression free

survival are still seen across all four CMS subtypes. Because stroma-derived mRNAs in whole tumour transcriptomes may obscure transcriptional features displayed by cancer cells, other efforts leveraged the power of patient-derived mouse xenograft (PDX) models in which human stroma is replaced by mouse stroma to obtain five CRC 'intrinsic' (CRIS) molecular subtypes, termed CRIS-A to E[13]. CRIS-C was identified as a subtype associated with EGFR signalling and increased sensitivity to anti-EGFR therapy. However responses to anti-EGFR therapy strongly varied among the other four CRIS subtypes [13].

EGFR activation results in the activation of several downstream signalling pathways, including the PI3K/AKT and MAPK pathways[14]. The activation status of these key signalling pathways influences a variety of biological processes such as proliferation, apoptosis, cell migration, bioenergetics, immune responses, and angiogenesis. A different approach to investigate responses to anti-EGFR therapy is to determine the activation status of key signalling branches activated by EGFR receptors and their downstream effectors, supported by statistical or deterministic modelling[15]. Because processes such as proliferation and apoptosis are controlled by complex networks that show significant signalling redundancies, deterministic systems models have been developed to estimate more precisely proliferative capacity or apoptosis sensitivity of tumours. One such tool developed by our group is the systems model, DR_MOMP, which calculates the apoptosis sensitivity of tumours based on a quantitative analysis of BCL-2 family proteins and their interactions[16, 17]. To identify novel prognostic markers of anti-EGFR therapy, we here

comprehensively profiled 83 signalling proteins and (phospho)proteins related to EGFR and key cancer signalling pathways in a cohort of 63 ~~'quadruple negative' (~~*KRAS*, *BRAF*, *NRAS* and *PI3KCA* wild type~~)~~ PDX models isolated from liver biopsies that were derived from metastatic CRC patients[9, 18]. We performed both statistical and systems modelling analyses to identify novel protein signatures of anti-EGFR responsiveness.

Methods

CRC PDX in vivo model

108 PDX models derived from colorectal cancer liver metastasis originally at the Institute for Cancer Research and Treatment, and Mauriziano Umberto I (Torino, Italy)[18] were used in this study. 63 of 108 were *KRAS*, *BRAF*, *NRAS* and *PI3KCA* wild type ~~quadruple negative (with wild-type KRas, NRas, PI3KCA, and B-Raf)~~ based on matched next-generation sequencing analysis data from Bertotti *et al.* [9] and used for statistical analysis. Tumour tissues were implanted subcutaneously and passaged in *NOD*/*SCID* mice. Response data is available for each tumour to cetuximab treatment after 3 and 6 weeks[19].

Reverse phase protein array

Protein was extracted from PDX tumour tissue and cell line standards and RPPA was performed as described previously[20]. Protein lysates normalized to 1µg/µL concentration as assessed by bicinchoninic acid assay (BCA, Biorad). Reverse phase protein array (RPPA) with a panel of antibodies targeting various key cancer related proteins was used for

measuring protein levels in untreated tumours. The response is form matching samples of same tumour in different mice. The DAKO (Carpinteria, CA) catalyzed signal amplification system was used for antibody blotting.

PDX Protein clustering

RPPA data for 93 PDX samples have been clustered using consensus Non-negative Matrix Factorization (R package 'NMF'[21], version 0.21.0) on centred RPPA data[22, 23]. NMF was performed 1 000 times with the number of clusters k varying from 2 to 8. k = 3 was selected based on visual inspection of co-clustering matrices and heatmap of clustered RPPA data.

To represent graphically the correspondence between CRIS subtypes classifiers and the RPPA clusters or cetuximab response, Factorial Correspondence Analysis (FCA) was used. For each comparison, $\chi 2$ independence test was carried out. In order to have large enough numbers in the contingency table so that the $\chi 2$ approximation is correct, we combined together the closest CRIS subtypes.

DR MOMP, APOPTO-CELL and proliferation signature

The normalised gene expression of *BIRC5*, *CCNB1*, *CDC20*, *CDCA1*, *CEP55*, *NDC80*, *MKI67*, *PTTG1*, *RRM2*, *TYMS* and *UBE2C* was averaged and used as proliferation signature[24, 25] of each PDX. The gene expression data for respective PDX models was downloaded from GSE76402[13].

To calculate the sensitivity of patients' cancer cells to undergo apoptosis, the mathematical models APOPTO-CELL[26] and DR_MOMP[16] were applied, using PRO-CASPASE-3, PRO-CASPASE-9, SMAC, and XIAP

protein for APOPTO-CELL, and BAK, BAX, BCL2 and BCL(X)L for DR_MOMP as input for the models. MCL1 protein levels were assumed to be 0 nM for DR_MOMP. SMAC concentrations were assumed to be 122.7 nM for APOPTO-CELL[26]. Protein levels were normalized to HeLa cells that were placed on the RPPA together with the cancer tissue[16, 26].

Statistical analysis

Statistical analysis of RPPA data was done using 'SAMR'[27] (Significance Analysis of Microarrays, version 3.0) and 'PAMR'[28] (Prediction Analysis for Microarrays; version 1.56.1) R Packages (R version 3.6.2). LASSO was performed using the 'glmnet' R package (version 2.0-18). The packages 'ComplexHeatmap'[29] (version 2.1.0) and 'Circlize'[30] (version 0.4.7) were used to create Figure 1. Week 3 response was used for all the statistical analysis as not all the mice were followed through after 3 weeks. Student's t-test and ANOVA was used for measuring statistical significance. ANOVA was followed by Tukey's HSD (honest significant difference) test for multiple pair comparison. Fisher's exact test was used for count data.

Results

Characterisation of *KRAS*, *BRAF*, *NRAS* and *PI3KCA* wild type metastatic CRC (phospho)protein signatures

To investigate cetuximab responses in patients with metastatic CRC, we analyzed a large collection of genomically annotated PDX models, for which information on response to cetuximab in mice was available[18]. Of the 108 patient-derived xenografts (PDX) *'KRAS* wild-type' models

originally collected (determined by Sanger sequencing), 63 samples were identified to bear no somatic sequence alteration of the *KRAS*, *NRAS*, *BRAF* and *PIK3CA* genes as identified by exome sequencing with an average coverage within the target regions of nearly 150-fold for each sample [9]. Protein levels were quantitatively profiled by Reverse Phase Protein Array (RPPA) analysis of fresh-frozen pre-treatment tumour samples derived from each PDX model (Figure 1A; Supplementary Table 1).

To explore whether cetuximab responses were related to differences in cell signalling pathways as evaluated by RPPA (phospho)protein analysis, we first performed unsupervised clustering using Nonnegative Matrix Factorizations (NMF) of the 63 quadruple negative samples (Supplementary Table 2). Clustering identified three distinct protein clusters termed C1, C2 and C3 (Figure 1A). We also performed clustering in all n = 93 *KRAS* wild type samples and found 88.9% consistency of the clusters (Supplementary Table 1).

Protein cluster C1 contained 35 PDX models of which 13 were regressing, 14 showed no change in volume, and 8 were progressing at week 3 (Figure 1B). Samples in C1 had predominantly high levels of phosphorylated Chk-1 (S345), c-RAF (S338), S6 Ribosomal Protein (S235/236 and S240/244), Gab-1 (Y627) and GSK-3β (S9; Figure 1A and Supplementary Figure 1). In contrast, C1 samples had low levels of phosphorylated p38 MAPK (T180/Y182), AMPK (T172), FAK (Y925), Src (Y527), and Src (Y416). Furthermore, samples had low levels of SMAC, BCL(X) and STAT3 proteins.

Cluster C2 contained 18 PDXs of which 4 were regressing, 10 showed no change in volume, and 4 were progressing after cetuximab treatment (Figure 1B). C2 tissues were characterised by high levels of phosphorylated EGFR (Y1068), BCL2 (S70 and T56), Src (Y527), and STAT3 (Y705) (Figure 1A and Supplementary Figure 1). Furthermore, the cluster had low p27 and PTEN levels. This cluster was also characterised by low levels of phosphorylated GSK-3β (S9), MAPK (T202/Y204) and MEK1/2 (S217/221).

Interestingly, cluster C3 contained no progressing tumour models, 6 with no change in volume and 4 regressing PDX models (Figure 1B). C3 tissues had high levels of phosphorylated p38 MAPK (T180/Y182), AKT (S473), MEK1/2 (S217/221), MAPK (T202/Y204) and PDK1 (S241), together with high levels of p70 S6 Kinase and p27 protein levels (Figure 1A and Supplementary Figure 1). Compared to clusters C1 and C2, C3 showed low IGFI-Rβ, PARP, cIAP-1, APAF-1 and EGFR protein levels, together with low levels of cleaved caspase 9 (D330).

There was no difference in genetic alterations between the clusters (not shown). Overall, *TP53* mutations were found in 90% (n = 57; from 89% in C1 to 94% in C2), *APC* mutations in 89% (n = 56; from 89% in C1 to 90% in C3) and *TTN* mutations in 48% (n = 30; from 40% in C1 to 70% in C3) of PDX models (genetic data from Bertotti *et al.*[9]). Further, we did not find protein clusters to be significantly associated with a specific CRIS molecular subtype (Figure 1C). C1 consisted of 4 CRIS-A, 7 CRIS-B, 16 CRIS- C, 5 CRIS-D and 3 CRIS- E. C2 consisted of 2 CRIS-A, 3 CRIS-B, 16 CRIS-C, 3 CRIS-D and 4 CRIS-E. C3 consisted of zero CRIS-A, 1

CRIS-B, 7 CRIS-C, 1 CRIS-D and 1 CRIS-E. Likely due to the small size of the tested collection, we did not find significant differences in response relative to the CRIS subtypes (Fisher's exact p = 0.49; Figure 1D).

Identification of a (phospho)protein signature predicting responses to cetuximab

In a subsequent analysis we used a statistical method for class prediction from gene expression data using nearest shrunken centroids (prediction analysis for microarrays; PAM)[28] to determine to what extent proteins were either up- or down-regulated in all PDX models when grouped according to their response to cetuximab at week 3 (Figure 2; Supplementary Table 3). Overall, proteins levels were found to be inverted when comparing regressing models with progressing models. Progressing tumour models had high levels of phosphorylated EGFR (Y1173 and Y1068), AKT (S373), S6 ribosomal protein (S235/236 and S240/244), HER3 (Y1289), cRAF (S338), Gab-1 (Y627) and BCL2 (T56), together with high protein levels of cIAP-1, IGFI-Rβ, PARP, BAK, BAX, EGFR and APAF-1 compared to regressing models. In contrast, levels of phosphorylated PDK1 (S241), Shc (Y317), STAT3 (Y705), FAK (Y925), phosphorylated GSK-3β (S9), Src (Y416), MAPK (T202/Y204), NF-κB-p65 (S536), Caspase-8, p27, Src, Xiap and SMAC were low in progressing compared to regressing models. When comparing responses at week 6, we observed high levels of AKT (S473), HIAP-2 and PARP, and low p27 levels in progressing compared to regressing models (Supplementary Figure 2).

Refinement of a (phospho)protein response score

As a next step, we aimed to further reduce the number of proteins required for a predictive (phospho)protein signature. For this purpose we employed least absolute shrinkage and selection operator (LASSO; L1 regularization) and binominal logistic regression (progression *versus* regression) to identify the variables strongest associated with treatment response from the markers identified above. The advantage of LASSO is that the method exploits sparsity by shrinking less important features' coefficients to zero. Using only progressing (n = 12) or regressing (n = 22) PDX models, LASSO reduced the required proteins to 22 markers (Figure 3AB): PDK1 (S241; β = 2.4687), Caspase-8 (β = 2.3486), Shc (Y317; β = 0.2415), Stat3 (Y705; β = 1.4916), p27 (β = 1.5234), XIAP (β = 0.2372), GSK-3β (S9; β = 1.3425), PI3-Kinase p110α (β = 0.4648), HER3 (β = 0.2071), cleaved Caspase-9 (D330; β = 0.0043), MAPK - ERK 1/2 (β = 0.2350) and PKC-alpha (S657; β = 0.9340) were found with a positive coefficient (Figure 3B). BAK (β = -1.6263), EGFR (Y1068; β = -0.1290), Akt (S473; β = -2.5973), S6 Ribosomal Protein (S240/244; β = -1.6658), HER3 (Y1289; β = -1.9349), mTOR (β = -1.600), NF-κB-p65 (S536; β = -1.9424), Gab-1 (Y627; β = -1.5928) and Bcl-2 (T56; β = -0.5066) were found with a negative coefficient (Figure 3B). The interception was 2.2000. To gain a deeper understanding of the role of these markers, we used the Spearman correlation coefficients (Figure 3A) to construct a co-expression network (Figure 3B). While proteins such as EGFR (Y1068) and NF-κB-p65 (S536) had the same coefficient in the LASSO model and were co-expressed, Shc (Y317), GSK-3β (S9), HER3, Caspase-8, PDK1 (S241),

BAK and mTOR had disagreeing signs. Assuming that co-expressed proteins fell in the same, active or respectively inactive, signalling pathway and hence conducted a similar signal, the disagreement in the coefficients' sign suggested a critical difference of the proteins' role in responses to cetuximab.

We then applied the regression model to the PDXs that showed no or only minor changes in tumour volume (n = 30), in order to test whether the model is able to define models with any increase in tumour volume as "progressing" (n = 16) or "regressing" (n = 14). Although this is a challenging task, the model identified 12 models as true "progressing" (true positive), 9 as true "regressing" (true negative), 5 "regressing" as "progressing" and 4 "progressing" as "regressing" models. Hence the majority of marginally progressing or regressing PDXs were correctly identified by the regression model.

Comparison of pre- and post-treatment protein profiles

In further exploratory analysis, we also investigated whether cetuximab treatment altered protein levels during treatment. We randomly selected 15 PDX models, one from protein cluster C1, seven from cluster C2 and seven from cluster C3. Protein quantification using RPPA were repeated for pre- and post-treatment tumour tissues on a separate RPPA run. The pre-treated PDX tissues had a mean correlation coefficient of 0.79 (25^{th} - 75^{th} percentile = 0.74 – 0.85) compared with the post-treated tissues (Supplementary Figure 3). Pairwise comparison of pre- and post-treatment samples showed that 6 out of 69 (phospho)proteins were significantly altered by more (or less) than factor 2 (or ½) in response to cetuximab.

Levels of phosphorylated Gab-1 (Y627; $p < 0.001$), MEK1/2 (S217/221; $p < 0.001$), p70 S6 kinase (T389; $p < 0.001$) and GSK-3β (S9; $p < 0.01$), together with levels of MEK1 ($p < 0.001$), cleaved Caspase-7 (D198; $p < 0.1$) proteins, were significantly lower in post-treatment compared to pre-treatment tissues (Figure 3C). The full list of changes in protein levels can be seen in Supplementary Table 4.

Levels of only 2 of the 6 proteins that were differential expressed were prognostic for the response to cetuximab when measured prior to treatment. Models not responding to cetuximab were more likely to lack Gab-1 (Y627) and GSK-3β (S9; Figure 2). Abundance of MEK1/2 (S217/221) was characteristic for models of the protein cluster without progressing tumours (C3, Supplementary Figure 1). Levels of p70 S6 kinase (T389; $p < 0.001$), MEK1 ($p < 0.001$) and cleaved Caspase-7 (D198; $p < 0.1$) were neither associated with a specific response to cetuximab nor a protein cluster.

Proliferation rather than apoptosis systems score predicts responses to cetuximab

To determine whether apoptosis competence was a prognostic marker for anti-EGFR therapy responses, we used protein levels of BCL-2, BCL-XL, MCL-1, BAX, BAK, APAF1, SMAC, XIAP, PROCASPASE-3 and -9 in the 63 PDX models as model inputs for two deterministic models of apoptosis competence, one describing the process of mitochondrial permeabilization, DR_MOMP[16], and one the process of caspase activation downstream of mitochondrial permeabilization, APOPTO-CELL[26] (Figure

4A). Both models were developed and validated by our group and previously shown to be prognostic for survival of stage 2 and 3 CRC patients[16, 17, 31]. DR_MOMP calculates the 'stress dose' of tumour cells required to undergo mitochondrial permeabilisation, with low values indicating a high apoptosis competence[16]. For quantitative evaluation of protein levels, cell lysates of the PDX models were normalized to lysates of HeLa cells in which absolute protein levels were previously determined by quantitative Western blotting using purified proteins[16, 26]. The mean levels of the proteins required as model inputs are shown in Figure 4BC. Employing DR_MOMP using the generated quantitative protein profiles, we determined a mean 'stress dose' of 171.4 nM (SD 56.4 nM) across all PDXs. PDXs with a 'stress dose' greater than the mean also had significantly less cleaved caspase 9 (D330) compared to models with 'stress dose' less than the mean (t-test $p < 0.01$), confirming impaired apoptosis in models with high DR_MOMP 'stress dose' values. However, the DR_MOMP score did not correlate with cetuximab responses (ANOVA $p = 0.6$; Figure 4E). The DR_MOMP apoptosis score was lowest in PDX models in cluster C1 (mean = 152.9 nM) and, greatest in C3 (mean = 246.0 nM; ANOVA $p < 0.0001$, Tukey post-hoc $p \leq 0.02$; Figure 4F). There were no significant differences in DR_MOMP apoptosis scores when PDXs were grouped based on the CRIS subtypes (ANOVA $p = 0.6$; Figure 4G).

APOPTO-CELL predicts apoptosis susceptibility of cells by modelling activation of executioner caspases and cleavage of their downstream substrates[26]. Exceeding a threshold of 25% substrate cleavage within 300

minutes served in previous studies as a surrogate for the competence of cells to undergo executioner (caspase 3) activation, in line with previous single-cell imaging findings [26, 31]. APOPTO-CELL identified 24 PDX samples with less than 25% predicted substrate cleavage and 36 models with more than 25 % predicted substrate cleavage. However the predicted substrate cleavage did not correlate with responses of the PDX models to cetuximab (Fisher's exact p = 0.89; Figure 4E). Further, there was no significant difference in the number of PDXs with substrate cleavage less or greater than 25% between protein clusters C1-C3 (Fisher's exact p = 0.09) or CRIS subtypes (Fisher's exact p = 0.85; Figures 4FG).

We also questioned whether apoptosis signalling contributed to cetuximab responses only in specific protein clusters/molecular subtypes. There was no significant differences between DR_MOMP 'stress dose' scores and treatment responses when PDX models broken down into the three protein clusters C1, C2 and C3 (ANOVA interaction p = 0.9) or into the CRIS subtypes (ANOVA interaction p = 0.9). Similarly, there was no significant differences between the APOPTO-CELL class and treatment responses after stratifying for the protein cluster or CRIS (not-adjusted Fisher's exact p > 0.12). Collectively, these data suggest that BCL2-dependent mitochondrial apoptosis and caspase-3 activation does not play a major role in cetuximab responses.

Next, we calculated the individual proliferative capacity of each PDX using an 11 gene signature index[24, 25] using existing gene expression profiles[13]. Numerically, proliferation indices were lowest in protein cluster C3, and highest in C2. Statistical analysis revealed no significantly differences

between protein clusters (ANOVA p = 0.1; Figure 4H). CRIS-D had significant higher indices compared to the CRIS-B molecular subtype (Tukey post-hoc p = 0.02) and C (Tukey post-hoc p < 0.001; ANOVA p = 0.001). Across all PDXs, the proliferation index gradually increased from PDXs with regressing toward progressing responses to cetuximab (ANOVA p-value of 0.01, Figure 4 J). Progressing PDX models had higher proliferation indices compared to stable (Tukey post-hoc p = 0.01 and 0.03) or regressing PDX models (Tukey post-hoc p = 0.001 and 0.02) if adjusted for either CRIS (ANOVA p = 0.01) or protein clusters (ANOVA p = 0.02). Collectively, these data suggested that proliferation rather than apoptosis score is a key determinant of cetuximab responses in 'quadruple negative' metastatic CRC PDX models.

~~We also condensed the cell death scores of DR_MOMP and APOPTO-CELL and the proliferation score to an overall growth score by classifying models with impaired apoptosis and high proliferation as high growth (n = 19), models with impeccable apoptosis competency and low proliferation as low growth (n = 6), and all other models as intermediate growth (n = 35; Figure 3K). Growths score did not reflect response to cetuximab with the PDX models being equally likely to show progression or regression in response to cetuximab (Fisher's Exact p = 0.18; Figure 3L).~~

Development of an improved (phospho)protein response score

Because our previous protein analysis identified cell death markers (Figure 2 and 3B; BAK, BCL2, cleaved Caspase-9, XIAP, etc.) that indicated responses to cetuximab, we finally decided to repeat the LASSO analysis with the 22 proteins, but replaced the apoptosis-related markers (BAK, BCL-2 (T56), cleaved Caspase-9 (D330) and XIAP) with the normalised DR_MOMP score. In addition, we removed the protein markers for AKT, mTOR, MAPK-ERK1/2 and PI3-Kinase p110α based on the assumption that these markers will likely not indicate the activation status of their respective signalling pathway. This enabled us to reduce the overall number of proteins analysed. The LASSO analysis set only the coefficient of DR_MOMP to zero: PDK1 (S241; β = 6.3505), Caspase-8 (β = 5.2772), Shc (Y317; β = 4.2598), Stat3 (Y705; β = 2.6455), p27 (β = 0.6169), GSK-3β (S9; β = 6.0001), HER3 (β = 3.5702) and PKC-alpha (S657; β = 0.8191) were found with a positive coefficient. EGFR (Y1068; β = -1.065), Akt (S473; β = -5.5777), S6 Ribosomal Protein (S240/244; β = -4.3452), HER3 (Y1289; β = -5.4732), NF-kB-p65 (S536; β = -6.3106) and Gab-1 (Y627; β = -4.6551) were found with a negative coefficients. The interception was 4.9424. The coefficients were in line with the first LASSO model (Spearman's rank correlation rho = 0.88, $p < 0.0001$). Testing the updated regression model (14 markers) on PDX models showing no or only minor changes in tumour volume (n = 30), showed a significant improvement compared with the initial score, with 13 PDX models identified as true "progressing" (true positive), 10 as true "regressing" (true

negative), 4 “regressing” as “progressing” and 3 ”progressing” as “regressing” models.

Discussion

The discovery of new prognostic biomarkers for cetuximab response is of crucial importance for improving efficiency, and efficacy, of the treatment of metastatic CRC. The genetic heterogeneity of metastatic CRC cancer makes it unlikely that one single protein will serve as a biomarker in all instances, and high throughput techniques such as RPPA may therefore be helpful in identifying predictive biomarker sets. Statistical analysis of our RPPA data showed significant correlation between levels of 20 (phospho)proteins with changes in tumour volume, as detected in PDX models. We identified markers indicating active signalling of the EGFR pathway such as EGFR (Y1068) itself and Akt (S473), Gab-1 (Y627), Shc (Y317), Stat3 (Y705) and PDK1 (S241) to significantly predict responses to cetuximab. Overall we found a high cross correlation between levels of these proteins markers across all samples, emphasising their potential to act as predictive biomarkers for cetuximab responses.

Interestingly, we found that high levels of phosphorylated EGFR at Tyr1068 and Akt at Ser473 indicated tumour progression, whereas regressing tumours showed a lack of phosphorylated Shc at Tyr317 and Stat3 at Tyr705. Phosphorylation of EGFR on Tyr1068 (and Tyr1086) leads to activation of the MAPK cascade and AKT activation[32]. *Signal transducer and activator of transcription 3* (STAT3) and its phosphorylation are associated with cell growth and transformation[33]. The scaffolding

protein *Src homology and collagen domain protein* (Shc) directs the EGF stimuli to pro-mitogenic, pro-survival and invasion signalling pathways in a time-dependent manner[34]. *Phosphoinositide Dependent Protein Kinase 1* (PDK1) is a crucial enzyme in transducing signals to multiple effector pathways including *phosphoinositide 3-kinase* (PI3K/AKT), *Ras/mitogen-activated protein kinase* (MAPK), *serum/glucocorticoid regulated kinase* (SGK), *p70 ribosomal protein S6 kinase* (p70 S6 K) and members of *protein kinase C* (PKC) family. Phosphorylation of PDK1 on Ser241 is necessary for its activation[35]. Some of its substrates require a prior conformational switch to allow subsequent phosphorylation by PDK1[35] rendering it as gatekeeper for those signalling pathways. We also found that models expressing the *human epidermal growth factor receptor 3* (HER3, also EGFR3) were more likely to respond with tumour regression in response to cetuximab. In contrast, phosphorylation of HER3 on Tyr1289 was indicative for tumour progression. HER3 cannot be activated by ligand alone but its heterodimer with EGFR and HER2 is highly mitogenic[36]. Existing literature on the expression and relevance of *HER3* is inconsistent, reporting association with either increased or decreased survival of CRC patients[36]. In advanced non-small cell lung cancer, abundant *HER3* expression identifies gefitinib (EGFR inhibitor) sensitive cell lines[37]. In addition, Bosch-Vilaró *et al.*[38] described a cetuximab-induced feedback HER3 activation that reduces the response to cetuximab, and in pancreatic cancer, dimerization of EGFR and HER3 was reported to be necessary for downstream signalling[39].

Further LASSO and binominal logistic regression analysis of these protein biomarkers delivered a refined protein signatures for predicting responses to cetuximab. Given that many of the identified markers in our signature are predicted to regulate cell proliferation, we also investigated a previously published, transcriptome-based proliferation score as to its predictive power[24, 25]. Using this score, we also found a significant correlation between cetuximab responses and the transcriptome-based proliferation score across all 63 PDX models investigated. Although the focus of our study was the delivery of a (phospo)protein signature, combining our protein score with the transcriptome-based proliferation score did not further increase the predictive power of the protein signature, suggesting that the signature was sufficient to describe the proliferation status of the PDX models in relation to cetuximab responses.

We also found that responses to cetuximab were dependent on protein clusters identified through unsupervised cluster analysis. One of the clusters, protein cluster 3 (C3), represented a cluster without progressing PDX models. C3 was characterised by PDK1-dependend active AKT signalling and inhibition of the cell cycle. The largest protein cluster (C1) in contrast showed mixed responses, and was characterised by genotoxic stress, inflammation and cell survival signalling. Cluster C2 was also composed of mixed responders and characterised by active EGFR signalling and inhibition of apoptosis. Compared to PDX models in C1 and C2, PDX models in C3 had lower levels of phosphorylated MEK1/2 (S217/221). This suggests that cetuximab-resistant models in C1 and C2 may potentially benefit from MEK inhibitors. We also explored the

relationship between protein clusters and transcriptome-based molecular subtypes. CRIS molecular subtypes capture very well differences in intrinsic tumour cell gene expression[13]. CRIS-C was previously associated with sensitivity to cetuximab[13], potentially a consequence of the lower representation of *KRAS* and *NRAS* mutations in this subtype[13]. We did not find that any of the three protein clusters showed a significant association with CRIS molecular subtypes. We also found that, when focusing on *KRAS*, *NRAS*, *BRAF* and *PIK3CA* ~~quadruple~~ wild type models, CRIS-C was not enriched in cetuximab responders (Figure 1D). Overall, this suggests that sensitivity to anti-EGFR therapy is predicted well by an analysis of (phospho)protein clusters.

While we observed that increased proliferative capacity was associated with disease progression during cetuximab treatment (Figure 4J), competence to undergo mitochondrial apopotosis was not a major determinant of cetuximab responses. Both the DR_MOMP and APOPTO-CELL apoptosis models have been shown to be prognostic for stage II and III CRC patients, but have not yet not been tested in the setting of metastatic CRC[17, 31].

Our data suggest that resistance to mitochondrial apoptosis is not critical for responses of metastatic CRC to cetuximab. While cetuximab was shown to induce apoptosis to a minor extent in colorectal cancer cells in previous studies[40], combination therapy for example with regorafenib has been shown to be required for significant apoptosis induction by cetuximab[41]. In the setting of colorectal cancer, we have previously also shown that activation of Caspase-3 may be associated with a

compensatory stimulation of cancer cell proliferation and adverse effects on clinical outcome[42]. Here, we also observed that PDX models with progressing tumours tended to have higher levels cleaved Caspase-3 compared to models with stable or regressing tumours (Figure 2). It might be possible that activating apoptosis may have both beneficial and detrimental effects in the setting of metastatic CRC.

By comparing matched pre- and post-treatment samples, we also found that levels of GSK-3β (S9) were reduced in tissue after cetuximab treatment. The *Glycogen synthase kinase 3*β (GSK-3β) is a key player in the β-catenin/Wnt signalling pathway but also phosphorylates various transcription factors and structural, metabolic and signalling proteins[43, 44]. Inhibition of GSK-3β activity by phosphorylation at Ser9[45] is a critical factor to allow many coupled signalling pathways to proceed[43, 44]. 96% of CRCs harbour increased oncogenic Wnt pathway alteration[46] and dysregulation of GSK-3β signalling is associated with cancer and metabolic and degenerative disorders[47]. Inhibition of GSK-3β was reported to induce apoptosis and attenuated proliferation in colon cancer cells *in vitro*[48] and in colon cancer xenografts[49]. It is possible that inhibition of GSK-3β would be desirable co-treatment with cetuximab. Lithium, which also acts as an inhibitor of GSK-3β[50], was reported to supress cell proliferation in prostate cancer xenographs[51] and may inhibit colon cancer metastatsis by blocking *transforming growth factor-β-induced* protein (TGFBIp) expression[52] downstream of GSK-3[53]. Combining cetuximab with lithium or other GSK-3β inhibitors may improve response to cetuximab.

In conclusion, we present here a 14 (phospho)protein marker signature that was predicting responses to cetuximab in mCRC tissue. Likewise, our findings emphasises GSK-3β to be potentially targetable for a co-treatment with cetuximab.

Further Disclosures

Ethics approval and consent to participate: Informed consent for research use was obtained from all patients and the study was conducted under the approval of the RCSI Research Ethics Committee and *Comitato Etico Istituto di Candiolo*-FPO IRCCS. All animal procedures were approved by the Ethical Commission of the Candiolo Cancer Institute and by the Italian Ministry of Health (806/2016-PR).

Data Accessibility: Data is provided as supplementary materials. Extended data and scripts will be made available upon reasonable request.

Conflict of interest: The authors declare no conflict of interest.

Funding: This study was supported by grants from Science Foundation Ireland and the Health Research Board to JHMP (13/IA/1881, 14/IA/2582, 15/ERACSM/3268 and 16/US/3301). LT is supported by AIRC (Associazione Italiana per la Ricerca sul Cancro) Investigator Grant 22802, AIRC 5x1000 grant 21091 (to LT), AIRC/CRUK/FC AECC Accelerator Award 22795, Fondazione Piemontese per la Ricerca sul Cancro-ONLUS, and 5x1000 Ministero della Salute 2014, 2015 and 2016. LT is a member of the EuroPDX Consortium.

Authors' contributions: ABa, AUL, SC, NM, MS and JHMP wrote the manuscript. ABa, AUL, SC, NM and MS performed data analysis and prepared figures. ABe and ERZ performed acquisition of sample data. BTH, ERZ, ROB, SC and MC collected samples and conducted the protein quantification using RPPA. ABe, BTH, LT and JHMP supervised the project. All authors read, reviewed and approved the final manuscript for publication.

References

1. Van Cutsem E, Cervantes A, Adam R, Sobrero A, Van Krieken JH, Aderka D, Aranda Aguilar E, Bardelli A, Benson A, Bodoky G, Ciardiello F, D'Hoore A, et al. ESMO consensus guidelines for the management of patients with metastatic colorectal cancer. *Ann Oncol* 2016;**27**: 1386-422.
2. Goldberg RM, Sargent DJ, Morton RF, Fuchs CS, Ramanathan RK, Williamson SK, Findlay BP, Pitot HC, Alberts SR. A randomized controlled trial of fluorouracil plus leucovorin, irinotecan, and oxaliplatin combinations in patients with previously untreated metastatic colorectal cancer. *J Clin Oncol* 2004;**22**: 23-30.
3. Van Cutsem E, Kohne CH, Hitre E, Zaluski J, Chang Chien CR, Makhson A, D'Haens G, Pinter T, Lim R, Bodoky G, Roh JK, Folprecht G, et al. Cetuximab and chemotherapy as initial treatment for metastatic colorectal cancer. *N Engl J Med* 2009;**360**: 1408-17.
4. Sepulveda AR, Hamilton SR, Allegra CJ, Grody W, Cushman-Vokoun AM, Funkhouser WK, Kopetz SE, Lieu C, Lindor NM, Minsky BD, Monzon FA, Sargent DJ, et al. Molecular Biomarkers for the Evaluation of Colorectal Cancer: Guideline From the American Society for Clinical Pathology, College of American Pathologists, Association for Molecular Pathology, and the American Society of Clinical Oncology. *J Clin Oncol* 2017;**35**: 1453-86.
5. Chiorean EG, Nandakumar G, Fadelu T, Temin S, Alarcon-Rozas AE, Bejarano S, Croitoru AE, Grover S, Lohar PV, Odhiambo A, Park SH, Garcia ER, et al. Treatment of Patients With Late-Stage Colorectal Cancer: ASCO Resource-Stratified Guideline. *JCO Glob Oncol* 2020;**6**: 414-38.
6. De Roock W, Claes B, Bernasconi D, De Schutter J, Biesmans B, Fountzilas G, Kalogeras KT, Kotoula V, Papamichael D, Laurent-Puig P, Penault-Llorca F, Rougier P, et al. Effects of KRAS, BRAF, NRAS, and PIK3CA mutations on the efficacy of cetuximab plus chemotherapy in chemotherapy-refractory metastatic colorectal cancer: a retrospective consortium analysis. *Lancet Oncol* 2010;**11**: 753-62.
7. De Roock W, De Vriendt V, Normanno N, Ciardiello F, Tejpar S. KRAS, BRAF, PIK3CA, and PTEN mutations: implications for targeted therapies in metastatic colorectal cancer. *Lancet Oncol* 2011;**12**: 594-603.
8. Zanella ER, Galimi F, Sassi F, Migliardi G, Cottino F, Leto SM, Lupo B, Erriquez J, Isella C, Comoglio PM, Medico E, Tejpar S, et al. IGF2 is an actionable target that identifies a distinct subpopulation of colorectal cancer patients with marginal response to anti-EGFR therapies. *Sci Transl Med* 2015;**7**: 272ra12.
9. Bertotti A, Papp E, Jones S, Adleff V, Anagnostou V, Lupo B, Sausen M, Phallen J, Hruban CA, Tokheim C, Niknafs N, Nesselbush M, et al. The genomic landscape of response to EGFR blockade in colorectal cancer. *Nature* 2015;**526**: 263-7.
10. Woolston A, Khan K, Spain G, Barber LJ, Griffiths B, Gonzalez-Exposito R, Hornsteiner L, Punta M, Patil Y, Newey A, Mansukhani S, Davies MN, et al. Genomic and Transcriptomic Determinants of Therapy Resistance and Immune Landscape Evolution during Anti-EGFR Treatment in Colorectal Cancer. *Cancer Cell* 2019;**36**: 35-50 e9.
11. Guinney J, Dienstmann R, Wang X, de Reynies A, Schlicker A, Soneson C, Marisa L, Roepman P, Nyamundanda G, Angelino P, Bot BM, Morris JS, et al. The consensus molecular subtypes of colorectal cancer. *Nat Med* 2015;**21**: 1350-6.
12. Lenz HJ, Ou FS, Venook AP, Hochster HS, Niedzwiecki D, Goldberg RM, Mayer RJ, Bertagnolli MM, Blanke CD, Zemla T, Qu X, Wirapati P, et al. Impact of Consensus Molecular Subtype on Survival in Patients With Metastatic Colorectal Cancer: Results From CALGB/SWOG 80405 (Alliance). *J Clin Oncol* 2019;**37**: 1876-85.
13. Isella C, Brundu F, Bellomo SE, Galimi F, Zanella E, Porporato R, Petti C, Fiori A, Orzan F, Senetta R, Boccaccio C, Ficarra E, et al. Selective analysis of cancer-cell intrinsic transcriptional traits defines novel clinically relevant subtypes of colorectal cancer. *Nat Commun* 2017;**8**: 15107.

14. Scaltriti M, Baselga J. The epidermal growth factor receptor pathway: a model for targeted therapy. *Clin Cancer Res* 2006;**12**: 5268-72.
15. Marzi L, Combes E, Vie N, Ayrolles-Torro A, Tosi D, Desigaud D, Perez-Gracia E, Larbouret C, Montagut C, Iglesias M, Jarlier M, Denis V, et al. FOXO3a and the MAPK p38 are activated by cetuximab to induce cell death and inhibit cell proliferation and their expression predicts cetuximab efficacy in colorectal cancer. *Br J Cancer* 2016;**115**: 1223-33.
16. Lindner AU, Concannon CG, Boukes GJ, Cannon MD, Llambi F, Ryan D, Boland K, Kehoe J, McNamara DA, Murray F, Kay EW, Hector S, et al. Systems analysis of BCL2 protein family interactions establishes a model to predict responses to chemotherapy. *Cancer Res* 2013;**73**: 519-28.
17. Lindner AU, Salvucci M, Morgan C, Monsefi N, Resler AJ, Cremona M, Curry S, Toomey S, O'Byrne R, Bacon O, Stuhler M, Flanagan L, et al. BCL-2 system analysis identifies high-risk colorectal cancer patients. *Gut* 2017;**66**: 2141-8.
18. Bertotti A, Migliardi G, Galimi F, Sassi F, Torti D, Isella C, Cora D, Di Nicolantonio F, Buscarino M, Petti C, Ribero D, Russolillo N, et al. A molecularly annotated platform of patient-derived xenografts ("xenopatients") identifies HER2 as an effective therapeutic target in cetuximab-resistant colorectal cancer. *Cancer Discov* 2011;**1**: 508-23.
19. Puig I, Chicote I, Tenbaum SP, Arques O, Herance JR, Gispert JD, Jimenez J, Landolfi S, Caci K, Allende H, Mendizabal L, Moreno D, et al. A personalized preclinical model to evaluate the metastatic potential of patient-derived colon cancer initiating cells. *Clin Cancer Res* 2013;**19**: 6787-801.
20. Guo H, Liu W, Ju Z, Tamboli P, Jonasch E, Mills GB, Lu Y, Hennessy BT, Tsavachidou D. An efficient procedure for protein extraction from formalin-fixed, paraffin-embedded tissues for reverse phase protein arrays. *Proteome Sci* 2012;**10**: 56.
21. Gaujoux R, Seoighe C. A flexible R package for nonnegative matrix factorization. *BMC Bioinformatics* 2010;**11**: 367.
22. Lee DD, Seung HS. Learning the parts of objects by non-negative matrix factorization. *Nature* 1999;**401**: 788-91.
23. Brunet JP, Tamayo P, Golub TR, Mesirov JP. Metagenes and molecular pattern discovery using matrix factorization. *Proc Natl Acad Sci U S A* 2004;**101**: 4164-9.
24. Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, Davies S, Fauron C, He X, Hu Z, Quackenbush JF, Stijleman IJ, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol* 2009;**27**: 1160-7.
25. Nielsen TO, Parker JS, Leung S, Voduc D, Ebbert M, Vickery T, Davies SR, Snider J, Stijleman IJ, Reed J, Cheang MC, Mardis ER, et al. A comparison of PAM50 intrinsic subtyping with immunohistochemistry and clinical prognostic factors in tamoxifen-treated estrogen receptor-positive breast cancer. *Clin Cancer Res* 2010;**16**: 5222-32.
26. Rehm M, Huber HJ, Dussmann H, Prehn JH. Systems analysis of effector caspase activation and its control by X-linked inhibitor of apoptosis protein. *EMBO J* 2006;**25**: 4338-49.
27. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 2001;**98**: 5116-21.
28. Tibshirani R, Hastie T, Narasimhan B, Chu G. Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc Natl Acad Sci U S A* 2002;**99**: 6567-72.
29. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 2016;**32**: 2847-9.
30. Gu Z, Gu L, Eils R, Schlesner M, Brors B. circlize Implements and enhances circular visualization in R. *Bioinformatics* 2014;**30**: 2811-2.
31. Salvucci M, Wurstle ML, Morgan C, Curry S, Cremona M, Lindner AU, Bacon O, Resler AJ, Murphy AC, O'Byrne R, Flanagan L, Dasgupta S, et al. A Stepwise Integrated Approach to Personalized Risk Predictions in Stage III Colorectal Cancer. *Clin Cancer Res* 2017;**23**: 1200-12.

32. Rodrigues GA, Falasca M, Zhang Z, Ong SH, Schlessinger J. A novel positive feedback loop mediated by the docking protein Gab1 and phosphatidylinositol 3-kinase in epidermal growth factor receptor signaling. *Mol Cell Biol* 2000;**20**: 1448-59.
33. Bromberg J, Darnell JE, Jr. The role of STATs in transcriptional control and their impact on cellular function. *Oncogene* 2000;**19**: 2468-73.
34. Zheng Y, Zhang C, Croucher DR, Soliman MA, St-Denis N, Pasculescu A, Taylor L, Tate SA, Hardy WR, Colwill K, Dai AY, Bagshaw R, et al. Temporal regulation of EGF signalling networks by the scaffold protein Shc1. *Nature* 2013;**499**: 166-71.
35. Casamayor A, Morrice NA, Alessi DR. Phosphorylation of Ser-241 is essential for the activity of 3-phosphoinositide-dependent protein kinase-1: identification of five sites of phosphorylation in vivo. *Biochem J* 1999;**342 (Pt 2)**: 287-92.
36. Campbell MR, Amin D, Moasser MM. HER3 comes of age: new insights into its functions and role in signaling, tumor biology, and cancer therapy. *Clin Cancer Res* 2010;**16**: 1373-83.
37. Engelman JA, Janne PA, Mermel C, Pearlberg J, Mukohara T, Fleet C, Cichowski K, Johnson BE, Cantley LC. ErbB-3 mediates phosphoinositide 3-kinase activity in gefitinib-sensitive non-small cell lung cancer cell lines. *Proc Natl Acad Sci U S A* 2005;**102**: 3788-93.
38. Bosch-Vilaro A, Jacobs B, Pomella V, Abbasi Asbagh L, Kirkland R, Michel J, Singh S, Liu X, Kim P, Weitsman G, Barber PR, Vojnovic B, et al. Feedback activation of HER3 attenuates response to EGFR inhibitors in colon cancer cells. *Oncotarget* 2017;**8**: 4277-88.
39. Frolov A, Schuller K, Tzeng CW, Cannon EE, Ku BC, Howard JH, Vickers SM, Heslin MJ, Buchsbaum DJ, Arnoletti JP. ErbB3 expression and dimerization with EGFR influence pancreatic cancer cell sensitivity to erlotinib. *Cancer Biol Ther* 2007;**6**: 548-54.
40. Li X, Fan Z. The epidermal growth factor receptor antibody cetuximab induces autophagy in cancer cells by downregulating HIF-1alpha and Bcl-2 and activating the beclin 1/hVps34 complex. *Cancer Res* 2010;**70**: 5942-52.
41. Napolitano S, Martini G, Rinaldi B, Martinelli E, Donniacuo M, Berrino L, Vitagliano D, Morgillo F, Barra G, De Palma R, Merolla F, Ciardiello F, et al. Primary and Acquired Resistance of Colorectal Cancer to Anti-EGFR Monoclonal Antibody Can Be Overcome by Combined Treatment of Regorafenib with Cetuximab. *Clin Cancer Res* 2015;**21**: 2975-83.
42. Flanagan L, Meyer M, Fay J, Curry S, Bacon O, Duessmann H, John K, Boland KC, McNamara DA, Kay EW, Bantel H, Schulze-Bergkamen H, et al. Low levels of Caspase-3 predict favourable response to 5FU-based chemotherapy in advanced colorectal cancer: Caspase-3 inhibition as a therapeutic approach. *Cell Death Dis* 2016;**7**: e2087.
43. Grimes CA, Jope RS. The multifaceted roles of glycogen synthase kinase 3beta in cellular signaling. *Prog Neurobiol* 2001;**65**: 391-426.
44. Woodgett JR. Judging a protein by more than its name: GSK-3. *Sci STKE* 2001;**2001**: re12.
45. Frame S, Cohen P, Biondi RM. A common phosphate binding site explains the unique substrate specificity of GSK3 and its inactivation by phosphorylation. *Mol Cell* 2001;**7**: 1321-7.
46. Yaeger R, Chatila WK, Lipsyc MD, Hechtman JF, Cercek A, Sanchez-Vega F, Jayakumaran G, Middha S, Zehir A, Donoghue MTA, You D, Viale A, et al. Clinical Sequencing Defines the Genomic Landscape of Metastatic Colorectal Cancer. *Cancer Cell* 2018;**33**: 125-36 e3.
47. Martinez A. Preclinical efficacy on GSK-3 inhibitors: towards a future generation of powerful drugs. *Med Res Rev* 2008;**28**: 773-96.
48. Shakoori A, Ougolkov A, Yu ZW, Zhang B, Modarressi MH, Billadeau DD, Mai M, Takahashi Y, Minamoto T. Deregulated GSK3beta activity in colorectal cancer: its association with tumor cell survival and proliferation. *Biochem Biophys Res Commun* 2005;**334**: 1365-73.
49. Shakoori A, Mai W, Miyashita K, Yasumoto K, Takahashi Y, Ooi A, Kawakami K, Minamoto T. Inhibition of GSK-3 beta activity attenuates proliferation of human colon cancer cells in rodents. *Cancer Sci* 2007;**98**: 1388-93.
50. Phiel CJ, Klein PS. Molecular targets of lithium action. *Annu Rev Pharmacol Toxicol* 2001;**41**: 789-813.

51. Zhu Q, Yang J, Han S, Liu J, Holzbeierlein J, Thrasher JB, Li B. Suppression of glycogen synthase kinase 3 activity reduces tumor growth of prostate cancer in vivo. *Prostate* 2011;**71**: 835-45.
52. Maeng YS, Lee R, Lee B, Choi SI, Kim EK. Lithium inhibits tumor lymphangiogenesis and metastasis through the inhibition of TGFBIp expression in cancer cells. *Sci Rep* 2016;**6**: 20739.
53. Liang MH, Chuang DM. Differential roles of glycogen synthase kinase-3 isoforms in the regulation of transcriptional activation. *J Biol Chem* 2006;**281**: 30479-84.

Figure Legends

Figure 1

(**A**) Heatmap of protein levels determined by RPPA. PDX models were annotated with the, CRIS, the consensus protein cluster subtype, and response to cetuximab (top). Clustering was performed using Nonnegative Matrix Factorization (NMF) consensus clustering algorithm. The right annotations indicates proteins' association to the protein clusters (Supplementary Figure 1). Chord diagrams show overlap between RPPA clusters and (**B**) response to cetuximab and (**C**) CRIS, and (**D**) overlap between CRIS and response to cetuximab.

Figure 2

Protein scores indicating proteins' association to the PDX models' response to cetuximab. Proteins' scores for response to cetuximab after 3 week was calculated using PAM [27].

Figure 3

(A) Heatmap of Spearman's rank correlation coefficients for proteins associated with differences in response to cetuximab from Figure 2. (**B**) Undirected graph of proteins found to be relevant in LASSO analysis. Intensity and colour of the edges indicate the correlation coefficient of (A). Grouping based on the signs of the

correlation coefficients and signs of the coefficients found by LASSO are indicated black & white nodes and plus & minus icons, respectively. (**C**) Protein found to be differential expressed in PDX models after treatment with cetuximab, based on pairwise comparison and Benjamin & Hochberg adjusted p-value. Dashed red lines indicate 0.05 significance threshold for p-value, and 2-fold or 1/2-fold protein level. The protein marker names and n-fold differences (treated to un-treated) in brackets were added for proteins passing all thresholds.

Figure 4

(**A**) Simplified illustration of the apoptotic signalling modelled in DR_MOMP and APOPTO-CELL. Absolute protein levels normalised to HeLa cells were measured using RPPA and used as input for (**B**) DR_MOMP and (**C**) APOPTO-CELL. Calculated DR MOMP values against (**D**) APOPTO-CELLs' calculated substrate cleavage class with (**E**) differences in response to cetuximab, (**F**) RPPA protein cluster C3 and (**G**) CRIS. Calculated proliferation against (**H**) protein clusters, (**I**) CRIS and (**J**) response to cetuximab. ~~(**K**) The proliferation score was combined with both models to a tumour growth score. (**L**) n-numbers of tumour growth score classes against response to cetuximab, RPPA protein cluster and CRIS.~~

Systems analysis of protein signatures predicting Cetuximab responses in *KRAS*, *NRAS*, *BRAF* and *PIK3CA* wild-type patient-derived xenografts models of metastatic colorectal cancer

Andreas U. Lindner[1], Steven Carberry[1], Naser Monsefi[1], Ana Barat[1], Manuela Salvucci[1], Robert O'Byrne[1], Eugenia R. Zanella[2], Mattia Cremona[3], Bryan T. Hennessy[3], Andrea Bertotti[2,4], Livio Trusolino[2,4] and Jochen H.M. Prehn[1]

[1]Department of Physiology and Medical Physics and Centre Systems Medicine, Royal College of Surgeons in Ireland, Dublin, Ireland; [2]Translational Cancer Medicine, Surgical Oncology, and Clinical Trials Coordination, Candiolo Cancer Institute Fondazione del Piemonte per l'Oncologia IRCCS, Turin, Italy; [3]Department of Medical Oncology, Beaumont Hospital, Royal College of Surgeons in Ireland, Dublin, Ireland; [4]Department of Oncology, University of Turin Medical School, Turin, Italy.

Corresponding Author: Prof. Jochen H.M. Prehn, Centre for Systems Medicine, and Department of Physiology and Medical Physics, Royal College of Surgeons in Ireland, 123 St. Stephen's Green, Dublin 2, Ireland. Tel.: +353 (1) 402 2255. Fax: + 353 (1) 402 2447. eMail: prehn@rcsi.ie

Running title: Systems analysis of cetuximab responses

Novelty and Impact: A large fraction of patients with metastatic colorectal cancer do not respond to anti-EGFR therapy despite KRAS wild type tumours. Statistical analysis of RPPA data of colorectal cancer KRAS, BRAF, NRAS and PI3KCA wild type PDX models revealed a 14 - 20 (phospho)protein signature that was predicting responses to cetuximab. Our findings furthermore emphasise GSK-3β to be potentially targetable for a co-treatment with cetuximab.

Keywords: anti-EGFR, metastatic colorectal cancer, molecular subtyping, reverse-phase protein array, deterministic modelling, apoptosis, proliferation

Abbreviations: 5-FU, fluorouracil; ANOVA, analysis of variance; CMS, consensus molecular subtypes; CRC, colorectal cancer; CRIS, CRC intrinsic subtype; EGF, epidermal growth factor; EGFR, EGF receptor; LASSO, least absolute shrinkage and selection operator; NMF, non-negative matrix factorization; P, p-value; PAM, Prediction Analysis for Microarrays; PDX, patient-derived mouse xenograft; RPPA, reverse phase protein array; SC, substrate cleavage.

Abstract

Antibodies targeting the human epidermal growth factor receptor (*EGFR*) are used for the treatment of *RAS* wild-type metastatic colorectal cancer. A significant proportion of patients remains unresponsive to this therapy. Here, we performed a reverse phase protein array-based (phospho)protein analysis of 63 *KRAS*, *NRAS*, *BRAF* and *PIK3CA* wild-type metastatic CRC tumours. Responses of tumours to anti-EGFR therapy with cetuximab were recorded in patient-derived xenograft (PDX) models. Unsupervised hierarchical clustering of pre-treatment tumour tissue identified three clusters, of which cluster C3 was exclusively composed of responders. Clusters C1 and C2 showed mixed responses. None of the three protein clusters showed a significant correlation with transcriptome-based subtypes. Analysis of protein signatures across all PDXs identified 14 markers that discriminated cetuximab-sensitive and -resistant tumours: PDK1 (S241), Caspase-8, Shc (Y317), Stat3 (Y705), p27, GSK-3β (S9), HER3, PKC-α (S657), EGFR (Y1068), Akt (S473), S6 Ribosomal Protein (S240/244), HER3 (Y1289), NF-κB-p65 (S536) and Gab-1 (Y627). Least absolute shrinkage and selection operator and binominal logistic regression analysis delivered refined protein signatures for predicting response to cetuximab. (Phospo-)protein analysis of matched pre- and post-treated models furthermore showed significant reduction of Gab-1 (Y627) and GSK-3β (S9) exclusively in responding models, suggesting novel targets for treatment.

Background

Colorectal cancer (CRC) is the third and second most commonly diagnosed cancer in males and females, and the second most common cause of cancer-related deaths in the developed world. In the advanced setting, CRC is routinely treated with fluorouracil (5-FU)-based chemotherapy. 30% of CRC patients present in the metastatic setting[1] where response rates to palliative 5-FU/oxaliplatin- or 5-FU/irinotecan-based chemotherapy range between 40-50%. Median overall survival remains poor at around 16-19 months[2]. Identifying the importance of epidermal growth factor (*EGF*) signalling for the survival of CRC cells resulted in the development of targeted therapies that neutralize the oncogenic activity of EGF receptors (*EGFR*). Anti-EGFR therapies have significantly improved survival in metastatic CRC patients[3]. Guidelines recommend to test for *KRAS*, *NRAS* and *BRAF* mutations as well as microsatellite instability status in CRC patients being considered for anti-EGFR therapy[4, 5] on the bases of the ineffectiveness of anti-EGFR therapy in patients with activating *KRAS*, *BRAF*, and *NRAS* mutations[6], and favourable responses to immune check point inhibitors in microsatellite instability-high patients[4]. While *PI3KCA* mutational analysis is not recommended yet[4], *PIK3CA* exon 20 mutations were linked with a worse outcome compared with wild-type status in patients with metastatic colorectal cancer[7]. Nevertheless, between 50–60% of patients will not benefit from anti-EGFR treatment even when these are *KRAS*, *BRAF*, *NRAS* and *PI3KCA* wild type[7].

Mutations and copy number alterations in genes encoding for other survival signaling proteins have been shown to contribute to anti-EGFR resistance. For example, *HER2*-amplification, *IGF2* overexpression or increased MET activity resulted in reduced responses to anti-EGFR therapy, as demonstrated in patient-derived xenograft (PDX) models of metastatic CRC and in patients[8, 9]. Analysis of the genomic and transcriptomic landscape of anti-EGFR resistance in PDX models and patients furthermore identified mutations in *EGFR*, *FGFR1*, *PDGFRA*, and *MAP2K1* or loss of *NF1* to contribute to anti-EGF resistance[9, 10].

While identification of patient-specific genome alterations provides a personalised diagnosis that provides insights into anti-EGFR therapy responses and may open opportunities for personalised therapies, interpretation of often multiple genomic alterations found in most patients is not always straightforward. Other efforts to identify responders and non-responders to anti-EGFR therapy have therefore focussed on the power of unsupervised molecular subtyping of tumours. An international meta-analysis and bioinformatics effort led to the identification of four distinct subtypes in CRC, termed 'Consensus Molecular Subtypes' (CMS1-CMS4)[11]. A recent study demonstrated that CMS2 patients benefitted more from anti-EGFR therapy than patients treated with anti-angiogenic therapy, while the opposite was the case in CMS1 patients[12]. However predictions of anti-EGFR therapy responses in CMS3 and CMS4 patients were not possible, and significant variability in overall and progression free survival are still seen across all four CMS subtypes. Because stroma-derived mRNAs in whole tumour transcriptomes may obscure

transcriptional features displayed by cancer cells, other efforts leveraged the power of patient-derived mouse xenograft (PDX) models in which human stroma is replaced by mouse stroma to obtain five CRC 'intrinsic' (CRIS) molecular subtypes, termed CRIS-A to E[13]. CRIS-C was identified as a subtype associated with EGFR signalling and increased sensitivity to anti-EGFR therapy. However responses to anti-EGFR therapy strongly varied among the other four CRIS subtypes[13].

EGFR activation results in the activation of several downstream signalling pathways, including the PI3K/AKT and MAPK pathways[14]. The activation status of these key signalling pathways influences a variety of biological processes such as proliferation, apoptosis, cell migration, bioenergetics, immune responses, and angiogenesis. A different approach to investigate responses to anti-EGFR therapy is to determine the activation status of key signalling branches activated by EGFR receptors and their downstream effectors, supported by statistical or deterministic modelling[15]. Because processes such as proliferation and apoptosis are controlled by complex networks that show significant signalling redundancies, deterministic systems models have been developed to estimate more precisely proliferative capacity or apoptosis sensitivity of tumours. One such tool developed by our group is the systems model, DR_MOMP, which calculates the apoptosis sensitivity of tumours based on a quantitative analysis of BCL-2 family proteins and their interactions[16, 17]. To identify novel prognostic markers of anti-EGFR therapy, we here comprehensively profiled 83 signalling proteins and (phospho)proteins related to EGFR and key cancer signalling pathways in a cohort of 63

KRAS, *BRAF*, *NRAS* and *PI3KCA* wild type PDX models isolated from liver biopsies that were derived from metastatic CRC patients[9, 18]. We performed both statistical and systems modelling analyses to identify novel protein signatures of anti-EGFR responsiveness.

Methods

CRC PDX in vivo model

108 PDX models derived from colorectal cancer liver metastasis originally at the Institute for Cancer Research and Treatment, and Mauriziano Umberto I (Torino, Italy)[18] were used in this study. 63 of 108 were *KRAS*, *BRAF*, *NRAS* and *PI3KCA* wild type based on matched next-generation sequencing analysis data from Bertotti *et al.* [9] and used for statistical analysis. Tumour tissues were implanted subcutaneously and passaged in *NOD*/*SCID* mice. Response data is available for each tumour to cetuximab treatment after 3 and 6 weeks[19].

Reverse phase protein array

Protein was extracted from PDX tumour tissue and cell line standards and RPPA was performed as described previously[20]. Protein lysates normalized to 1µg/µL concentration as assessed by bicinchoninic acid assay (BCA, Biorad). Reverse phase protein array (RPPA) with a panel of antibodies targeting various key cancer related proteins was used for measuring protein levels in untreated tumours. The response is form matching samples of same tumour in different mice. The DAKO

(Carpinteria, CA) catalyzed signal amplification system was used for antibody blotting.

PDX Protein clustering

RPPA data for 93 PDX samples have been clustered using consensus Non-negative Matrix Factorization (R package 'NMF'[21], version 0.21.0) on centred RPPA data[22, 23]. NMF was performed 1 000 times with the number of clusters k varying from 2 to 8. k = 3 was selected based on visual inspection of co-clustering matrices and heatmap of clustered RPPA data.

To represent graphically the correspondence between CRIS subtypes classifiers and the RPPA clusters or cetuximab response, Factorial Correspondence Analysis (FCA) was used. For each comparison, $\chi 2$ independence test was carried out. In order to have large enough numbers in the contingency table so that the $\chi 2$ approximation is correct, we combined together the closest CRIS subtypes.

DR MOMP, APOPTO-CELL and proliferation signature

The normalised gene expression of *BIRC5*, *CCNB1*, *CDC20*, *CDCA1*, *CEP55*, *NDC80*, *MKI67*, *PTTG1*, *RRM2*, *TYMS* and *UBE2C* was averaged and used as proliferation signature[24, 25] of each PDX. The gene expression data for respective PDX models was downloaded from GSE76402[13].

To calculate the sensitivity of patients' cancer cells to undergo apoptosis, the mathematical models APOPTO-CELL[26] and DR_MOMP[16] were applied, using PRO-CASPASE-3, PRO-CASPASE-9, SMAC, and XIAP protein for APOPTO-CELL, and BAK, BAX, BCL2 and BCL(X)L for DR_MOMP as input for the models. MCL1 protein levels were assumed to

be 0 nM for DR_MOMP. SMAC concentrations were assumed to be 122.7 nM for APOPTO-CELL[26]. Protein levels were normalized to HeLa cells that were placed on the RPPA together with the cancer tissue[16, 26].

Statistical analysis

Statistical analysis of RPPA data was done using 'SAMR'[27] (Significance Analysis of Microarrays, version 3.0) and 'PAMR'[28] (Prediction Analysis for Microarrays; version 1.56.1) R Packages (R version 3.6.2). LASSO was performed using the 'glmnet' R package (version 2.0-18). The packages 'ComplexHeatmap'[29] (version 2.1.0) and 'Circlize'[30] (version 0.4.7) were used to create Figure 1. Week 3 response was used for all the statistical analysis as not all the mice were followed through after 3 weeks. Student's t-test and ANOVA was used for measuring statistical significance. ANOVA was followed by Tukey's HSD (honest significant difference) test for multiple pair comparison. Fisher's exact test was used for count data.

Results

Characterisation of *KRAS*, *BRAF*, *NRAS* and *PI3KCA* wild type metastatic CRC (phospho)protein signatures

To investigate cetuximab responses in patients with metastatic CRC, we analyzed a large collection of genomically annotated PDX models, for which information on response to cetuximab in mice was available[18]. Of the 108 patient-derived xenografts (PDX) *'KRAS* wild-type' models originally collected (determined by Sanger sequencing), 63 samples were identified to bear no somatic sequence alteration of the *KRAS*, *NRAS*,

BRAF and *PIK3CA* genes as identified by exome sequencing with an average coverage within the target regions of nearly 150-fold for each sample [9]. Protein levels were quantitatively profiled by Reverse Phase Protein Array (RPPA) analysis of fresh-frozen pre-treatment tumour samples derived from each PDX model (Figure 1A; Supplementary Table 1).

To explore whether cetuximab responses were related to differences in cell signalling pathways as evaluated by RPPA (phospho)protein analysis, we first performed unsupervised clustering using Nonnegative Matrix Factorizations (NMF) of the 63 quadruple negative samples (Supplementary Table 2). Clustering identified three distinct protein clusters termed C1, C2 and C3 (Figure 1A). We also performed clustering in all n = 93 *KRAS* wild type samples and found 88.9% consistency of the clusters (Supplementary Table 1).

Protein cluster C1 contained 35 PDX models of which 13 were regressing, 14 showed no change in volume, and 8 were progressing at week 3 (Figure 1B). Samples in C1 had predominantly high levels of phosphorylated Chk-1 (S345), c-RAF (S338), S6 Ribosomal Protein (S235/236 and S240/244), Gab-1 (Y627) and GSK-3β (S9; Figure 1A and Supplementary Figure 1). In contrast, C1 samples had low levels of phosphorylated p38 MAPK (T180/Y182), AMPK (T172), FAK (Y925), Src (Y527), and Src (Y416). Furthermore, samples had low levels of SMAC, BCL(X) and STAT3 proteins.

Cluster C2 contained 18 PDXs of which 4 were regressing, 10 showed no change in volume, and 4 were progressing after cetuximab treatment

(Figure 1B). C2 tissues were characterised by high levels of phosphorylated EGFR (Y1068), BCL2 (S70 and T56), Src (Y527), and STAT3 (Y705) (Figure 1A and Supplementary Figure 1). Furthermore, the cluster had low p27 and PTEN levels. This cluster was also characterised by low levels of phosphorylated GSK-3β (S9), MAPK (T202/Y204) and MEK1/2 (S217/221).

Interestingly, cluster C3 contained no progressing tumour models, 6 with no change in volume and 4 regressing PDX models (Figure 1B). C3 tissues had high levels of phosphorylated p38 MAPK (T180/Y182), AKT (S473), MEK1/2 (S217/221), MAPK (T202/Y204) and PDK1 (S241), together with high levels of p70 S6 Kinase and p27 protein levels (Figure 1A and Supplementary Figure 1). Compared to clusters C1 and C2, C3 showed low IGFI-Rβ, PARP, cIAP-1, APAF-1 and EGFR protein levels, together with low levels of cleaved caspase 9 (D330).

There was no difference in genetic alterations between the clusters (not shown). Overall, *TP53* mutations were found in 90% (n = 57; from 89% in C1 to 94% in C2), *APC* mutations in 89% (n = 56; from 89% in C1 to 90% in C3) and *TTN* mutations in 48% (n = 30; from 40% in C1 to 70% in C3) of PDX models (genetic data from Bertotti *et al.*[9]). Further, we did not find protein clusters to be significantly associated with a specific CRIS molecular subtype (Figure 1C). C1 consisted of 4 CRIS-A, 7 CRIS-B, 16 CRIS- C, 5 CRIS-D and 3 CRIS- E. C2 consisted of 2 CRIS-A, 3 CRIS-B, 16 CRIS-C, 3 CRIS-D and 4 CRIS-E. C3 consisted of zero CRIS-A, 1 CRIS-B, 7 CRIS-C, 1 CRIS-D and 1 CRIS-E. Likely due to the small size

of the tested collection, we did not find significant differences in response relative to the CRIS subtypes (Fisher's exact p = 0.49; Figure 1D).

Identification of a (phospho)protein signature predicting responses to cetuximab

In a subsequent analysis we used a statistical method for class prediction from gene expression data using nearest shrunken centroids (prediction analysis for microarrays; PAM)[28] to determine to what extent proteins were either up- or down-regulated in all PDX models when grouped according to their response to cetuximab at week 3 (Figure 2; Supplementary Table 3). Overall, proteins levels were found to be inverted when comparing regressing models with progressing models. Progressing tumour models had high levels of phosphorylated EGFR (Y1173 and Y1068), AKT (S373), S6 ribosomal protein (S235/236 and S240/244), HER3 (Y1289), cRAF (S338), Gab-1 (Y627) and BCL2 (T56), together with high protein levels of cIAP-1, IGFI-Rβ, PARP, BAK, BAX, EGFR and APAF-1 compared to regressing models. In contrast, levels of phosphorylated PDK1 (S241), Shc (Y317), STAT3 (Y705), FAK (Y925), phosphorylated GSK-3β (S9), Src (Y416), MAPK (T202/Y204), NF-κB-p65 (S536), Caspase-8, p27, Src, Xiap and SMAC were low in progressing compared to regressing models. When comparing responses at week 6, we observed high levels of AKT (S473), HIAP-2 and PARP, and low p27 levels in progressing compared to regressing models (Supplementary Figure 2).

Refinement of a (phospho)protein response score

As a next step, we aimed to further reduce the number of proteins required for a predictive (phospho)protein signature. For this purpose we employed least absolute shrinkage and selection operator (LASSO; L1 regularization) and binominal logistic regression (progression *versus* regression) to identify the variables strongest associated with treatment response from the markers identified above. The advantage of LASSO is that the method exploits sparsity by shrinking less important features' coefficients to zero. Using only progressing (n = 12) or regressing (n = 22) PDX models, LASSO reduced the required proteins to 22 markers (Figure 3AB): PDK1 (S241; β = 2.4687), Caspase-8 (β = 2.3486), Shc (Y317; β = 0.2415), Stat3 (Y705; β = 1.4916), p27 (β = 1.5234), XIAP (β = 0.2372), GSK-3β (S9; β = 1.3425), PI3-Kinase p110α (β = 0.4648), HER3 (β = 0.2071), cleaved Caspase-9 (D330; β = 0.0043), MAPK - ERK 1/2 (β = 0.2350) and PKC-alpha (S657; β = 0.9340) were found with a positive coefficient (Figure 3B). BAK (β = -1.6263), EGFR (Y1068; β = -0.1290), Akt (S473; β = -2.5973), S6 Ribosomal Protein (S240/244; β = -1.6658), HER3 (Y1289; β = -1.9349), mTOR (β = -1.600), NF-κB-p65 (S536; β = -1.9424), Gab-1 (Y627; β = -1.5928) and Bcl-2 (T56; β = -0.5066) were found with a negative coefficient (Figure 3B). The interception was 2.2000. To gain a deeper understanding of the role of these markers, we used the Spearman correlation coefficients (Figure 3A) to construct a co-expression network (Figure 3B). While proteins such as EGFR (Y1068) and NF-κB-p65 (S536) had the same coefficient in the LASSO model and were co-expressed, Shc (Y317), GSK-3β (S9), HER3, Caspase-8, PDK1 (S241),

BAK and mTOR had disagreeing signs. Assuming that co-expressed proteins fell in the same, active or respectively inactive, signalling pathway and hence conducted a similar signal, the disagreement in the coefficients' sign suggested a critical difference of the proteins' role in responses to cetuximab.

We then applied the regression model to the PDXs that showed no or only minor changes in tumour volume (n = 30), in order to test whether the model is able to define models with any increase in tumour volume as "progressing" (n = 16) or "regressing" (n = 14). Although this is a challenging task, the model identified 12 models as true "progressing" (true positive), 9 as true "regressing" (true negative), 5 "regressing" as "progressing" and 4 "progressing" as "regressing" models. Hence the majority of marginally progressing or regressing PDXs were correctly identified by the regression model.

Comparison of pre- and post-treatment protein profiles

In further exploratory analysis, we also investigated whether cetuximab treatment altered protein levels during treatment. We randomly selected 15 PDX models, one from protein cluster C1, seven from cluster C2 and seven from cluster C3. Protein quantification using RPPA were repeated for pre- and post-treatment tumour tissues on a separate RPPA run. The pre-treated PDX tissues had a mean correlation coefficient of 0.79 (25^{th} - 75^{th} percentile = 0.74 – 0.85) compared with the post-treated tissues (Supplementary Figure 3). Pairwise comparison of pre- and post-treatment samples showed that 6 out of 69 (phospho)proteins were significantly altered by more (or less) than factor 2 (or ½) in response to cetuximab.

Levels of phosphorylated Gab-1 (Y627; $p < 0.001$), MEK1/2 (S217/221; $p < 0.001$), p70 S6 kinase (T389; $p < 0.001$) and GSK-3β (S9; $p < 0.01$), together with levels of MEK1 ($p < 0.001$), cleaved Caspase-7 (D198; $p < 0.1$) proteins, were significantly lower in post-treatment compared to pre-treatment tissues (Figure 3C). The full list of changes in protein levels can be seen in Supplementary Table 4.

Levels of only 2 of the 6 proteins that were differential expressed were prognostic for the response to cetuximab when measured prior to treatment. Models not responding to cetuximab were more likely to lack Gab-1 (Y627) and GSK-3β (S9; Figure 2). Abundance of MEK1/2 (S217/221) was characteristic for models of the protein cluster without progressing tumours (C3, Supplementary Figure 1). Levels of p70 S6 kinase (T389; $p < 0.001$), MEK1 ($p < 0.001$) and cleaved Caspase-7 (D198; $p < 0.1$) were neither associated with a specific response to cetuximab nor a protein cluster.

Proliferation rather than apoptosis systems score predicts responses to cetuximab

To determine whether apoptosis competence was a prognostic marker for anti-EGFR therapy responses, we used protein levels of BCL-2, BCL-XL, MCL-1, BAX, BAK, APAF1, SMAC, XIAP, PROCASPASE-3 and -9 in the 63 PDX models as model inputs for two deterministic models of apoptosis competence, one describing the process of mitochondrial permeabilization, DR_MOMP[16], and one the process of caspase activation downstream of mitochondrial permeabilization, APOPTO-CELL[26] (Figure

4A). Both models were developed and validated by our group and previously shown to be prognostic for survival of stage 2 and 3 CRC patients[16, 17, 31]. DR_MOMP calculates the 'stress dose' of tumour cells required to undergo mitochondrial permeabilisation, with low values indicating a high apoptosis competence[16]. For quantitative evaluation of protein levels, cell lysates of the PDX models were normalized to lysates of HeLa cells in which absolute protein levels were previously determined by quantitative Western blotting using purified proteins[16, 26]. The mean levels of the proteins required as model inputs are shown in Figure 4BC. Employing DR_MOMP using the generated quantitative protein profiles, we determined a mean 'stress dose' of 171.4 nM (SD 56.4 nM) across all PDXs. PDXs with a 'stress dose' greater than the mean also had significantly less cleaved caspase 9 (D330) compared to models with 'stress dose' less than the mean (t-test $p < 0.01$), confirming impaired apoptosis in models with high DR_MOMP 'stress dose' values. However, the DR_MOMP score did not correlate with cetuximab responses (ANOVA $p = 0.6$; Figure 4E). The DR_MOMP apoptosis score was lowest in PDX models in cluster C1 (mean = 152.9 nM) and, greatest in C3 (mean = 246.0 nM; ANOVA $p < 0.0001$, Tukey post-hoc $p \leq 0.02$; Figure 4F). There were no significant differences in DR_MOMP apoptosis scores when PDXs were grouped based on the CRIS subtypes (ANOVA $p = 0.6$; Figure 4G).

APOPTO-CELL predicts apoptosis susceptibility of cells by modelling activation of executioner caspases and cleavage of their downstream substrates[26]. Exceeding a threshold of 25% substrate cleavage within 300

minutes served in previous studies as a surrogate for the competence of cells to undergo executioner (caspase 3) activation, in line with previous single-cell imaging findings [26, 31]. APOPTO-CELL identified 24 PDX samples with less than 25% predicted substrate cleavage and 36 models with more than 25 % predicted substrate cleavage. However the predicted substrate cleavage did not correlate with responses of the PDX models to cetuximab (Fisher's exact p = 0.89; Figure 4E). Further, there was no significant difference in the number of PDXs with substrate cleavage less or greater than 25% between protein clusters C1-C3 (Fisher's exact p = 0.09) or CRIS subtypes (Fisher's exact p = 0.85; Figures 4FG).

We also questioned whether apoptosis signalling contributed to cetuximab responses only in specific protein clusters/molecular subtypes. There was no significant differences between DR_MOMP 'stress dose' scores and treatment responses when PDX models broken down into the three protein clusters C1, C2 and C3 (ANOVA interaction p = 0.9) or into the CRIS subtypes (ANOVA interaction p = 0.9). Similarly, there was no significant differences between the APOPTO-CELL class and treatment responses after stratifying for the protein cluster or CRIS (not-adjusted Fisher's exact p > 0.12). Collectively, these data suggest that BCL2-dependent mitochondrial apoptosis and caspase-3 activation does not play a major role in cetuximab responses.

Next, we calculated the individual proliferative capacity of each PDX using an 11 gene signature index[24, 25] using existing gene expression profiles[13]. Numerically, proliferation indices were lowest in protein cluster C3, and highest in C2. Statistical analysis revealed no significantly differences

between protein clusters (ANOVA p = 0.1; Figure 4H). CRIS-D had significant higher indices compared to the CRIS-B molecular subtype (Tukey post-hoc p = 0.02) and C (Tukey post-hoc p < 0.001; ANOVA p = 0.001). Across all PDXs, the proliferation index gradually increased from PDXs with regressing toward progressing responses to cetuximab (ANOVA p-value of 0.01, Figure 4

J). Progressing PDX models had higher proliferation indices compared to stable (Tukey post-hoc p = 0.01 and 0.03) or regressing PDX models (Tukey post-hoc p = 0.001 and 0.02) if adjusted for either CRIS (ANOVA p = 0.01) or protein clusters (ANOVA p = 0.02). Collectively, these data suggested that proliferation rather than apoptosis score is a key determinant of cetuximab responses in 'quadruple negative' metastatic CRC PDX models.

Development of an improved (phospho)protein response score

Because our previous protein analysis identified cell death markers (Figure 2 and 3B; BAK, BCL2, cleaved Caspase-9, XIAP, etc.) that indicated responses to cetuximab, we finally decided to repeat the LASSO analysis with the 22 proteins, but replaced the apoptosis-related markers (BAK, BCL-2 (T56), cleaved Caspase-9 (D330) and XIAP) with the normalised DR_MOMP score. In addition, we removed the protein markers for AKT, mTOR, MAPK-ERK1/2 and PI3-Kinase p110α based on the assumption that these markers will likely not indicate the activation status of their respective signalling pathway. This enabled us to reduce the overall

number of proteins analysed. The LASSO analysis set only the coefficient of DR_MOMP to zero: PDK1 (S241; β = 6.3505), Caspase-8 (β = 5.2772), Shc (Y317; β = 4.2598), Stat3 (Y705; β = 2.6455), p27 (β = 0.6169), GSK-3β (S9; β = 6.0001), HER3 (β = 3.5702) and PKC-alpha (S657; β = 0.8191) were found with a positive coefficient. EGFR (Y1068; β = -1.065), Akt (S473; β = -5.5777), S6 Ribosomal Protein (S240/244; β = -4.3452), HER3 (Y1289; β = -5.4732), NF-kB-p65 (S536; β = -6.3106) and Gab-1 (Y627; β = -4.6551) were found with a negative coefficients. The interception was 4.9424. The coefficients were in line with the first LASSO model (Spearman's rank correlation rho = 0.88, $p < 0.0001$). Testing the updated regression model (14 markers) on PDX models showing no or only minor changes in tumour volume (n = 30), showed a significant improvement compared with the initial score, with 13 PDX models identified as true "progressing" (true positive), 10 as true "regressing" (true negative), 4 "regressing" as "progressing" and 3 "progressing" as "regressing" models.

Discussion

The discovery of new prognostic biomarkers for cetuximab response is of crucial importance for improving efficiency, and efficacy, of the treatment of metastatic CRC. The genetic heterogeneity of metastatic CRC cancer makes it unlikely that one single protein will serve as a biomarker in all instances, and high throughput techniques such as RPPA may therefore be helpful in identifying predictive biomarker sets. Statistical analysis of our RPPA data showed significant correlation between levels of 20

(phospho)proteins with changes in tumour volume, as detected in PDX models. We identified markers indicating active signalling of the EGFR pathway such as EGFR (Y1068) itself and Akt (S473), Gab-1 (Y627), Shc (Y317), Stat3 (Y705) and PDK1 (S241) to significantly predict responses to cetuximab. Overall we found a high cross correlation between levels of these proteins markers across all samples, emphasising their potential to act as predictive biomarkers for cetuximab responses.

Interestingly, we found that high levels of phosphorylated EGFR at Tyr1068 and Akt at Ser473 indicated tumour progression, whereas regressing tumours showed a lack of phosphorylated Shc at Tyr317 and Stat3 at Tyr705. Phosphorylation of EGFR on Tyr1068 (and Tyr1086) leads to activation of the MAPK cascade and AKT activation[32]. *Signal transducer and activator of transcription 3* (STAT3) and its phosphorylation are associated with cell growth and transformation[33]. The scaffolding protein *Src homology and collagen domain protein* (Shc) directs the EGF stimuli to pro-mitogenic, pro-survival and invasion signalling pathways in a time-dependent manner[34]. *Phosphoinositide Dependent Protein Kinase 1* (PDK1) is a crucial enzyme in transducing signals to multiple effector pathways including *phosphoinositide 3-kinase* (PI3K/AKT), *Ras/mitogen-activated protein kinase* (MAPK), *serum/glucocorticoid regulated kinase* (SGK), *p70 ribosomal protein S6 kinase* (p70 S6 K) and members of *protein kinase C* (PKC) family. Phosphorylation of PDK1 on Ser241 is necessary for its activation[35]. Some of its substrates require a prior conformational switch to allow subsequent phosphorylation by PDK1[35] rendering it as gatekeeper for those signalling pathways. We also found

that models expressing the *human epidermal growth factor receptor 3* (HER3, also EGFR3) were more likely to respond with tumour regression in response to cetuximab. In contrast, phosphorylation of HER3 on Tyr1289 was indicative for tumour progression. HER3 cannot be activated by ligand alone but its heterodimer with EGFR and HER2 is highly mitogenic[36]. Existing literature on the expression and relevance of *HER3* is inconsistent, reporting association with either increased or decreased survival of CRC patients[36]. In advanced non-small cell lung cancer, abundant *HER3* expression identifies gefitinib (EGFR inhibitor) sensitive cell lines[37]. In addition, Bosch-Vilaró *et al.*[38] described a cetuximab-induced feedback HER3 activation that reduces the response to cetuximab, and in pancreatic cancer, dimerization of EGFR and HER3 was reported to be necessary for downstream signalling[39].

Further LASSO and binominal logistic regression analysis of these protein biomarkers delivered a refined protein signatures for predicting responses to cetuximab. Given that many of the identified markers in our signature are predicted to regulate cell proliferation, we also investigated a previously published, transcriptome-based proliferation score as to its predictive power[24, 25]. Using this score, we also found a significant correlation between cetuximab responses and the transcriptome-based proliferation score across all 63 PDX models investigated. Although the focus of our study was the delivery of a (phospo)protein signature, combining our protein score with the transcriptome-based proliferation score did not further increase the predictive power of the protein

signature, suggesting that the signature was sufficient to describe the proliferation status of the PDX models in relation to cetuximab responses.

We also found that responses to cetuximab were dependent on protein clusters identified through unsupervised cluster analysis. One of the clusters, protein cluster 3 (C3), represented a cluster without progressing PDX models. C3 was characterised by PDK1-dependend active AKT signalling and inhibition of the cell cycle. The largest protein cluster (C1) in contrast showed mixed responses, and was characterised by genotoxic stress, inflammation and cell survival signalling. Cluster C2 was also composed of mixed responders and characterised by active EGFR signalling and inhibition of apoptosis. Compared to PDX models in C1 and C2, PDX models in C3 had lower levels of phosphorylated MEK1/2 (S217/221). This suggests that cetuximab-resistant models in C1 and C2 may potentially benefit from MEK inhibitors. We also explored the relationship between protein clusters and transcriptome-based molecular subtypes. CRIS molecular subtypes capture very well differences in intrinsic tumour cell gene expression[13]. CRIS-C was previously associated with sensitivity to cetuximab[13], potentially a consequence of the lower representation of *KRAS* and *NRAS* mutations in this subtype[13]. We did not find that any of the three protein clusters showed a significant association with CRIS molecular subtypes. We also found that, when focusing on *KRAS*, *NRAS*, *BRAF* and *PIK3CA* wild type models, CRIS-C was not enriched in cetuximab responders (Figure 1D). Overall, this suggests that sensitivity to anti-EGFR therapy is predicted well by an analysis of (phospho)protein clusters.

While we observed that increased proliferative capacity was associated with disease progression during cetuximab treatment (Figure 4J), competence to undergo mitochondrial apopotosis was not a major determinant of cetuximab responses. Both the DR_MOMP and APOPTO-CELL apoptosis models have been shown to be prognostic for stage II and III CRC patients, but have not yet not been tested in the setting of metastatic CRC[17, 31].

Our data suggest that resistance to mitochondrial apoptosis is not critical for responses of metastatic CRC to cetuximab. While cetuximab was shown to induce apoptosis to a minor extent in colorectal cancer cells in previous studies[40], combination therapy for example with regorafenib has been shown to be required for significant apoptosis induction by cetuximab[41]. In the setting of colorectal cancer, we have previously also shown that activation of Caspase-3 may be associated with a compensatory stimulation of cancer cell proliferation and adverse effects on clinical outcome[42]. Here, we also observed that PDX models with progressing tumours tended to have higher levels cleaved Caspase-3 compared to models with stable or regressing tumours (Figure 2). It might be possible that activating apoptosis may have both beneficial and detrimental effects in the setting of metastatic CRC.

By comparing matched pre- and post-treatment samples, we also found that levels of GSK-3β (S9) were reduced in tissue after cetuximab treatment. The *Glycogen synthase kinase 3*β (GSK-3β) is a key player in the β-catenin/Wnt signalling pathway but also phosphorylates various transcription factors and structural, metabolic and signalling proteins[43, 44].

Inhibition of GSK-3β activity by phosphorylation at Ser9[45] is a critical factor to allow many coupled signalling pathways to proceed[43, 44]. 96% of CRCs harbour increased oncogenic Wnt pathway alteration[46] and dysregulation of GSK-3β signalling is associated with cancer and metabolic and degenerative disorders[47]. Inhibition of GSK-3β was reported to induce apoptosis and attenuated proliferation in colon cancer cells *in vitro*[48] and in colon cancer xenografts[49]. It is possible that inhibition of GSK-3β would be desirable co-treatment with cetuxiumab. Lithium, which also acts as an inhibitor of GSK-3β[50], was reported to supress cell proliferation in prostate cancer xenographs[51] and may inhibit colon cancer metastatsis by blocking *transforming growth factor-β-induced* protein (TGFBIp) expression[52] downstream of GSK-3[53]. Combining cetuximab with lithium or other GSK-3β inhibitors may improve response to cetuximab.

In conclusion, we present here a 14 (phospho)protein marker signature that was predicting responses to cetuximab in mCRC tissue. Likewise, our findings emphasises GSK-3β to be potentially targetable for a co-treatment with cetuximab.

Further Disclosures

Ethics approval and consent to participate: Informed consent for research use was obtained from all patients and the study was conducted under the approval of the RCSI Research Ethics Committee and *Comitato Etico Istituto di Candiolo*-FPO IRCCS. All animal procedures were

approved by the Ethical Commission of the Candiolo Cancer Institute and by the Italian Ministry of Health (806/2016-PR).

Data Accessibility: Data is provided as supplementary materials. Extended data and scripts will be made available upon reasonable request.

Conflict of interest: The authors declare no conflict of interest.

Funding: This study was supported by grants from Science Foundation Ireland and the Health Research Board to JHMP (13/IA/1881, 14/IA/2582, 15/ERACSM/3268 and 16/US/3301). LT is supported by AIRC (Associazione Italiana per la Ricerca sul Cancro) Investigator Grant 22802, AIRC 5x1000 grant 21091 (to LT), AIRC/CRUK/FC AECC Accelerator Award 22795, Fondazione Piemontese per la Ricerca sul Cancro-ONLUS, and 5x1000 Ministero della Salute 2014, 2015 and 2016. LT is a member of the EuroPDX Consortium.

Authors' contributions: ABa, AUL, SC, NM, MS and JHMP wrote the manuscript. ABa, AUL, SC, NM and MS performed data analysis and prepared figures. ABe and ERZ performed acquisition of sample data. BTH, ERZ, ROB, SC and MC collected samples and conducted the protein quantification using RPPA. ABe, BTH, LT and JHMP supervised the project. All authors read, reviewed and approved the final manuscript for publication.

References

1. Van Cutsem E, Cervantes A, Adam R, Sobrero A, Van Krieken JH, Aderka D, Aranda Aguilar E, Bardelli A, Benson A, Bodoky G, Ciardiello F, D'Hoore A, et al. ESMO consensus guidelines for the management of patients with metastatic colorectal cancer. *Ann Oncol* 2016;**27**: 1386-422.
2. Goldberg RM, Sargent DJ, Morton RF, Fuchs CS, Ramanathan RK, Williamson SK, Findlay BP, Pitot HC, Alberts SR. A randomized controlled trial of fluorouracil plus leucovorin, irinotecan, and oxaliplatin combinations in patients with previously untreated metastatic colorectal cancer. *J Clin Oncol* 2004;**22**: 23-30.
3. Van Cutsem E, Kohne CH, Hitre E, Zaluski J, Chang Chien CR, Makhson A, D'Haens G, Pinter T, Lim R, Bodoky G, Roh JK, Folprecht G, et al. Cetuximab and chemotherapy as initial treatment for metastatic colorectal cancer. *N Engl J Med* 2009;**360**: 1408-17.
4. Sepulveda AR, Hamilton SR, Allegra CJ, Grody W, Cushman-Vokoun AM, Funkhouser WK, Kopetz SE, Lieu C, Lindor NM, Minsky BD, Monzon FA, Sargent DJ, et al. Molecular Biomarkers for the Evaluation of Colorectal Cancer: Guideline From the American Society for Clinical Pathology, College of American Pathologists, Association for Molecular Pathology, and the American Society of Clinical Oncology. *J Clin Oncol* 2017;**35**: 1453-86.
5. Chiorean EG, Nandakumar G, Fadelu T, Temin S, Alarcon-Rozas AE, Bejarano S, Croitoru AE, Grover S, Lohar PV, Odhiambo A, Park SH, Garcia ER, et al. Treatment of Patients With Late-Stage Colorectal Cancer: ASCO Resource-Stratified Guideline. *JCO Glob Oncol* 2020;**6**: 414-38.
6. De Roock W, Claes B, Bernasconi D, De Schutter J, Biesmans B, Fountzilas G, Kalogeras KT, Kotoula V, Papamichael D, Laurent-Puig P, Penault-Llorca F, Rougier P, et al. Effects of KRAS, BRAF, NRAS, and PIK3CA mutations on the efficacy of cetuximab plus chemotherapy in chemotherapy-refractory metastatic colorectal cancer: a retrospective consortium analysis. *Lancet Oncol* 2010;**11**: 753-62.
7. De Roock W, De Vriendt V, Normanno N, Ciardiello F, Tejpar S. KRAS, BRAF, PIK3CA, and PTEN mutations: implications for targeted therapies in metastatic colorectal cancer. *Lancet Oncol* 2011;**12**: 594-603.
8. Zanella ER, Galimi F, Sassi F, Migliardi G, Cottino F, Leto SM, Lupo B, Erriquez J, Isella C, Comoglio PM, Medico E, Tejpar S, et al. IGF2 is an actionable target that identifies a distinct subpopulation of colorectal cancer patients with marginal response to anti-EGFR therapies. *Sci Transl Med* 2015;**7**: 272ra12.
9. Bertotti A, Papp E, Jones S, Adleff V, Anagnostou V, Lupo B, Sausen M, Phallen J, Hruban CA, Tokheim C, Niknafs N, Nesselbush M, et al. The genomic landscape of response to EGFR blockade in colorectal cancer. *Nature* 2015;**526**: 263-7.
10. Woolston A, Khan K, Spain G, Barber LJ, Griffiths B, Gonzalez-Exposito R, Hornsteiner L, Punta M, Patil Y, Newey A, Mansukhani S, Davies MN, et al. Genomic and Transcriptomic Determinants of Therapy Resistance and Immune Landscape Evolution during Anti-EGFR Treatment in Colorectal Cancer. *Cancer Cell* 2019;**36**: 35-50 e9.
11. Guinney J, Dienstmann R, Wang X, de Reynies A, Schlicker A, Soneson C, Marisa L, Roepman P, Nyamundanda G, Angelino P, Bot BM, Morris JS, et al. The consensus molecular subtypes of colorectal cancer. *Nat Med* 2015;**21**: 1350-6.
12. Lenz HJ, Ou FS, Venook AP, Hochster HS, Niedzwiecki D, Goldberg RM, Mayer RJ, Bertagnolli MM, Blanke CD, Zemla T, Qu X, Wirapati P, et al. Impact of Consensus Molecular Subtype on Survival in Patients With Metastatic Colorectal Cancer: Results From CALGB/SWOG 80405 (Alliance). *J Clin Oncol* 2019;**37**: 1876-85.
13. Isella C, Brundu F, Bellomo SE, Galimi F, Zanella E, Porporato R, Petti C, Fiori A, Orzan F, Senetta R, Boccaccio C, Ficarra E, et al. Selective analysis of cancer-cell intrinsic transcriptional traits defines novel clinically relevant subtypes of colorectal cancer. *Nat Commun* 2017;**8**: 15107.

14. Scaltriti M, Baselga J. The epidermal growth factor receptor pathway: a model for targeted therapy. *Clin Cancer Res* 2006;**12**: 5268-72.

15. Marzi L, Combes E, Vie N, Ayrolles-Torro A, Tosi D, Desigaud D, Perez-Gracia E, Larbouret C, Montagut C, Iglesias M, Jarlier M, Denis V, et al. FOXO3a and the MAPK p38 are activated by cetuximab to induce cell death and inhibit cell proliferation and their expression predicts cetuximab efficacy in colorectal cancer. *Br J Cancer* 2016;**115**: 1223-33.

16. Lindner AU, Concannon CG, Boukes GJ, Cannon MD, Llambi F, Ryan D, Boland K, Kehoe J, McNamara DA, Murray F, Kay EW, Hector S, et al. Systems analysis of BCL2 protein family interactions establishes a model to predict responses to chemotherapy. *Cancer Res* 2013;**73**: 519-28.

17. Lindner AU, Salvucci M, Morgan C, Monsefi N, Resler AJ, Cremona M, Curry S, Toomey S, O'Byrne R, Bacon O, Stuhler M, Flanagan L, et al. BCL-2 system analysis identifies high-risk colorectal cancer patients. *Gut* 2017;**66**: 2141-8.

18. Bertotti A, Migliardi G, Galimi F, Sassi F, Torti D, Isella C, Cora D, Di Nicolantonio F, Buscarino M, Petti C, Ribero D, Russolillo N, et al. A molecularly annotated platform of patient-derived xenografts ("xenopatients") identifies HER2 as an effective therapeutic target in cetuximab-resistant colorectal cancer. *Cancer Discov* 2011;**1**: 508-23.

19. Puig I, Chicote I, Tenbaum SP, Arques O, Herance JR, Gispert JD, Jimenez J, Landolfi S, Caci K, Allende H, Mendizabal L, Moreno D, et al. A personalized preclinical model to evaluate the metastatic potential of patient-derived colon cancer initiating cells. *Clin Cancer Res* 2013;**19**: 6787-801.

20. Guo H, Liu W, Ju Z, Tamboli P, Jonasch E, Mills GB, Lu Y, Hennessy BT, Tsavachidou D. An efficient procedure for protein extraction from formalin-fixed, paraffin-embedded tissues for reverse phase protein arrays. *Proteome Sci* 2012;**10**: 56.

21. Gaujoux R, Seoighe C. A flexible R package for nonnegative matrix factorization. *BMC Bioinformatics* 2010;**11**: 367.

22. Lee DD, Seung HS. Learning the parts of objects by non-negative matrix factorization. *Nature* 1999;**401**: 788-91.

23. Brunet JP, Tamayo P, Golub TR, Mesirov JP. Metagenes and molecular pattern discovery using matrix factorization. *Proc Natl Acad Sci U S A* 2004;**101**: 4164-9.

24. Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, Davies S, Fauron C, He X, Hu Z, Quackenbush JF, Stijleman IJ, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol* 2009;**27**: 1160-7.

25. Nielsen TO, Parker JS, Leung S, Voduc D, Ebbert M, Vickery T, Davies SR, Snider J, Stijleman IJ, Reed J, Cheang MC, Mardis ER, et al. A comparison of PAM50 intrinsic subtyping with immunohistochemistry and clinical prognostic factors in tamoxifen-treated estrogen receptor-positive breast cancer. *Clin Cancer Res* 2010;**16**: 5222-32.

26. Rehm M, Huber HJ, Dussmann H, Prehn JH. Systems analysis of effector caspase activation and its control by X-linked inhibitor of apoptosis protein. *EMBO J* 2006;**25**: 4338-49.

27. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 2001;**98**: 5116-21.

28. Tibshirani R, Hastie T, Narasimhan B, Chu G. Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc Natl Acad Sci U S A* 2002;**99**: 6567-72.

29. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 2016;**32**: 2847-9.

30. Gu Z, Gu L, Eils R, Schlesner M, Brors B. circlize Implements and enhances circular visualization in R. *Bioinformatics* 2014;**30**: 2811-2.

31. Salvucci M, Wurstle ML, Morgan C, Curry S, Cremona M, Lindner AU, Bacon O, Resler AJ, Murphy AC, O'Byrne R, Flanagan L, Dasgupta S, et al. A Stepwise Integrated Approach to Personalized Risk Predictions in Stage III Colorectal Cancer. *Clin Cancer Res* 2017;**23**: 1200-12.

32. Rodrigues GA, Falasca M, Zhang Z, Ong SH, Schlessinger J. A novel positive feedback loop mediated by the docking protein Gab1 and phosphatidylinositol 3-kinase in epidermal growth factor receptor signaling. *Mol Cell Biol* 2000;**20**: 1448-59.
33. Bromberg J, Darnell JE, Jr. The role of STATs in transcriptional control and their impact on cellular function. *Oncogene* 2000;**19**: 2468-73.
34. Zheng Y, Zhang C, Croucher DR, Soliman MA, St-Denis N, Pasculescu A, Taylor L, Tate SA, Hardy WR, Colwill K, Dai AY, Bagshaw R, et al. Temporal regulation of EGF signalling networks by the scaffold protein Shc1. *Nature* 2013;**499**: 166-71.
35. Casamayor A, Morrice NA, Alessi DR. Phosphorylation of Ser-241 is essential for the activity of 3-phosphoinositide-dependent protein kinase-1: identification of five sites of phosphorylation in vivo. *Biochem J* 1999;**342 (Pt 2)**: 287-92.
36. Campbell MR, Amin D, Moasser MM. HER3 comes of age: new insights into its functions and role in signaling, tumor biology, and cancer therapy. *Clin Cancer Res* 2010;**16**: 1373-83.
37. Engelman JA, Janne PA, Mermel C, Pearlberg J, Mukohara T, Fleet C, Cichowski K, Johnson BE, Cantley LC. ErbB-3 mediates phosphoinositide 3-kinase activity in gefitinib-sensitive non-small cell lung cancer cell lines. *Proc Natl Acad Sci U S A* 2005;**102**: 3788-93.
38. Bosch-Vilaro A, Jacobs B, Pomella V, Abbasi Asbagh L, Kirkland R, Michel J, Singh S, Liu X, Kim P, Weitsman G, Barber PR, Vojnovic B, et al. Feedback activation of HER3 attenuates response to EGFR inhibitors in colon cancer cells. *Oncotarget* 2017;**8**: 4277-88.
39. Frolov A, Schuller K, Tzeng CW, Cannon EE, Ku BC, Howard JH, Vickers SM, Heslin MJ, Buchsbaum DJ, Arnoletti JP. ErbB3 expression and dimerization with EGFR influence pancreatic cancer cell sensitivity to erlotinib. *Cancer Biol Ther* 2007;**6**: 548-54.
40. Li X, Fan Z. The epidermal growth factor receptor antibody cetuximab induces autophagy in cancer cells by downregulating HIF-1alpha and Bcl-2 and activating the beclin 1/hVps34 complex. *Cancer Res* 2010;**70**: 5942-52.
41. Napolitano S, Martini G, Rinaldi B, Martinelli E, Donniacuo M, Berrino L, Vitagliano D, Morgillo F, Barra G, De Palma R, Merolla F, Ciardiello F, et al. Primary and Acquired Resistance of Colorectal Cancer to Anti-EGFR Monoclonal Antibody Can Be Overcome by Combined Treatment of Regorafenib with Cetuximab. *Clin Cancer Res* 2015;**21**: 2975-83.
42. Flanagan L, Meyer M, Fay J, Curry S, Bacon O, Duessmann H, John K, Boland KC, McNamara DA, Kay EW, Bantel H, Schulze-Bergkamen H, et al. Low levels of Caspase-3 predict favourable response to 5FU-based chemotherapy in advanced colorectal cancer: Caspase-3 inhibition as a therapeutic approach. *Cell Death Dis* 2016;**7**: e2087.
43. Grimes CA, Jope RS. The multifaceted roles of glycogen synthase kinase 3beta in cellular signaling. *Prog Neurobiol* 2001;**65**: 391-426.
44. Woodgett JR. Judging a protein by more than its name: GSK-3. *Sci STKE* 2001;**2001**: re12.
45. Frame S, Cohen P, Biondi RM. A common phosphate binding site explains the unique substrate specificity of GSK3 and its inactivation by phosphorylation. *Mol Cell* 2001;**7**: 1321-7.
46. Yaeger R, Chatila WK, Lipsyc MD, Hechtman JF, Cercek A, Sanchez-Vega F, Jayakumaran G, Middha S, Zehir A, Donoghue MTA, You D, Viale A, et al. Clinical Sequencing Defines the Genomic Landscape of Metastatic Colorectal Cancer. *Cancer Cell* 2018;**33**: 125-36 e3.
47. Martinez A. Preclinical efficacy on GSK-3 inhibitors: towards a future generation of powerful drugs. *Med Res Rev* 2008;**28**: 773-96.
48. Shakoori A, Ougolkov A, Yu ZW, Zhang B, Modarressi MH, Billadeau DD, Mai M, Takahashi Y, Minamoto T. Deregulated GSK3beta activity in colorectal cancer: its association with tumor cell survival and proliferation. *Biochem Biophys Res Commun* 2005;**334**: 1365-73.
49. Shakoori A, Mai W, Miyashita K, Yasumoto K, Takahashi Y, Ooi A, Kawakami K, Minamoto T. Inhibition of GSK-3 beta activity attenuates proliferation of human colon cancer cells in rodents. *Cancer Sci* 2007;**98**: 1388-93.
50. Phiel CJ, Klein PS. Molecular targets of lithium action. *Annu Rev Pharmacol Toxicol* 2001;**41**: 789-813.

51. Zhu Q, Yang J, Han S, Liu J, Holzbeierlein J, Thrasher JB, Li B. Suppression of glycogen synthase kinase 3 activity reduces tumor growth of prostate cancer in vivo. *Prostate* 2011;**71**: 835-45.
52. Maeng YS, Lee R, Lee B, Choi SI, Kim EK. Lithium inhibits tumor lymphangiogenesis and metastasis through the inhibition of TGFBIp expression in cancer cells. *Sci Rep* 2016;**6**: 20739.
53. Liang MH, Chuang DM. Differential roles of glycogen synthase kinase-3 isoforms in the regulation of transcriptional activation. *J Biol Chem* 2006;**281**: 30479-84.

Figure Legends

Figure 1

(**A**) Heatmap of protein levels determined by RPPA. PDX models were annotated with the, CRIS, the consensus protein cluster subtype, and response to cetuximab (top). Clustering was performed using Nonnegative Matrix Factorization (NMF) consensus clustering algorithm. The right annotations indicates proteins' association to the protein clusters (Supplementary Figure 1). Chord diagrams show overlap between RPPA clusters and (**B**) response to cetuximab and (**C**) CRIS, and (**D**) overlap between CRIS and response to cetuximab.

Figure 2

Protein scores indicating proteins' association to the PDX models' response to cetuximab. Proteins' scores for response to cetuximab after 3 week was calculated using PAM [27].

Figure 3

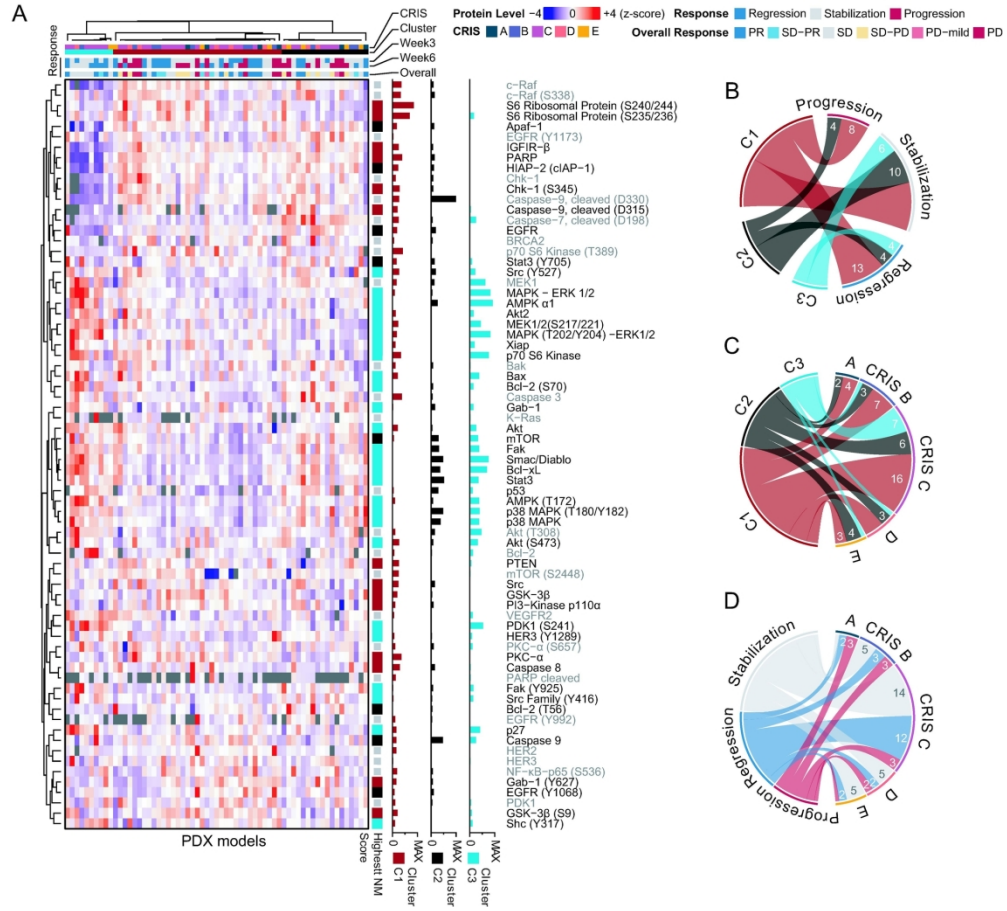
(A) Heatmap of Spearman's rank correlation coefficients for proteins associated with differences in response to cetuximab from Figure 2. (**B**) Undirected graph of proteins found to be relevant in LASSO analysis. Intensity and colour of the edges indicate the correlation coefficient of (A). Grouping based on the signs of the

correlation coefficients and signs of the coefficients found by LASSO are indicated black & white nodes and plus & minus icons, respectively. (**C**) Protein found to be differential expressed in PDX models after treatment with cetuximab, based on pairwise comparison and Benjamin & Hochberg adjusted p-value. Dashed red lines indicate 0.05 significance threshold for p-value, and 2-fold or 1/2-fold protein level. The protein marker names and n-fold differences (treated to un-treated) in brackets were added for proteins passing all thresholds.

Figure 4

(**A**) Simplified illustration of the apoptotic signalling modelled in DR_MOMP and APOPTO-CELL. Absolute protein levels normalised to HeLa cells were measured using RPPA and used as input for (**B**) DR_MOMP and (**C**) APOPTO-CELL. Calculated DR MOMP values against (**D**) APOPTO-CELLs' calculated substrate cleavage class with (**E**) differences in response to cetuximab, (**F**) RPPA protein cluster C3 and (**G**) CRIS. Calculated proliferation against (**H**) protein clusters, (**I**) CRIS and (**J**) response to cetuximab.

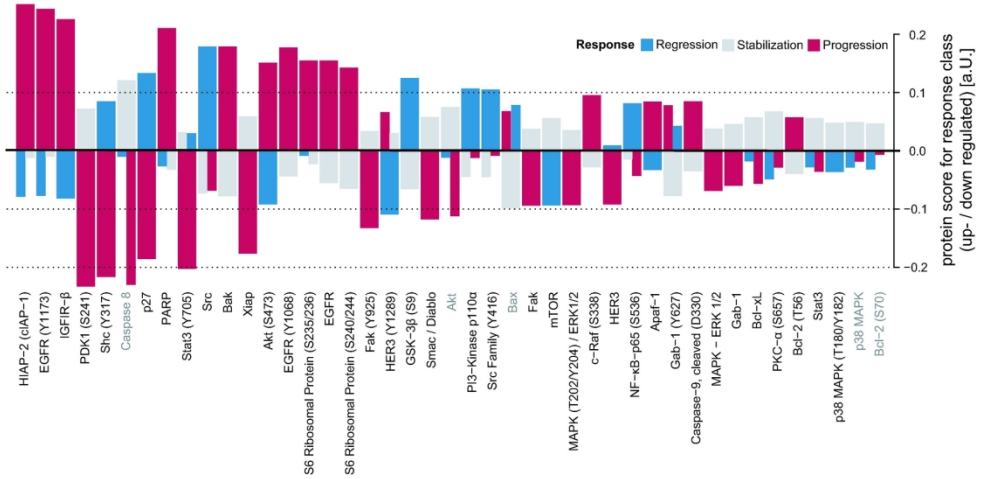
Figure 1



(A) Heatmap of protein levels determined by RPPA. PDX models were annotated with the, CRIS, the consensus protein cluster subtype, and response to cetuximab (top). Clustering was performed using Nonnegative Matrix Factorization (NMF) consensus clustering algorithm. The right annotations indicates proteins' association to the protein clusters (Supplementary Figure 1). Chord diagrams show overlap between RPPA clusters and (B) response to cetuximab and (C) CRIS, and (D) overlap between CRIS and response to cetuximab.

184x172mm (300 x 300 DPI)

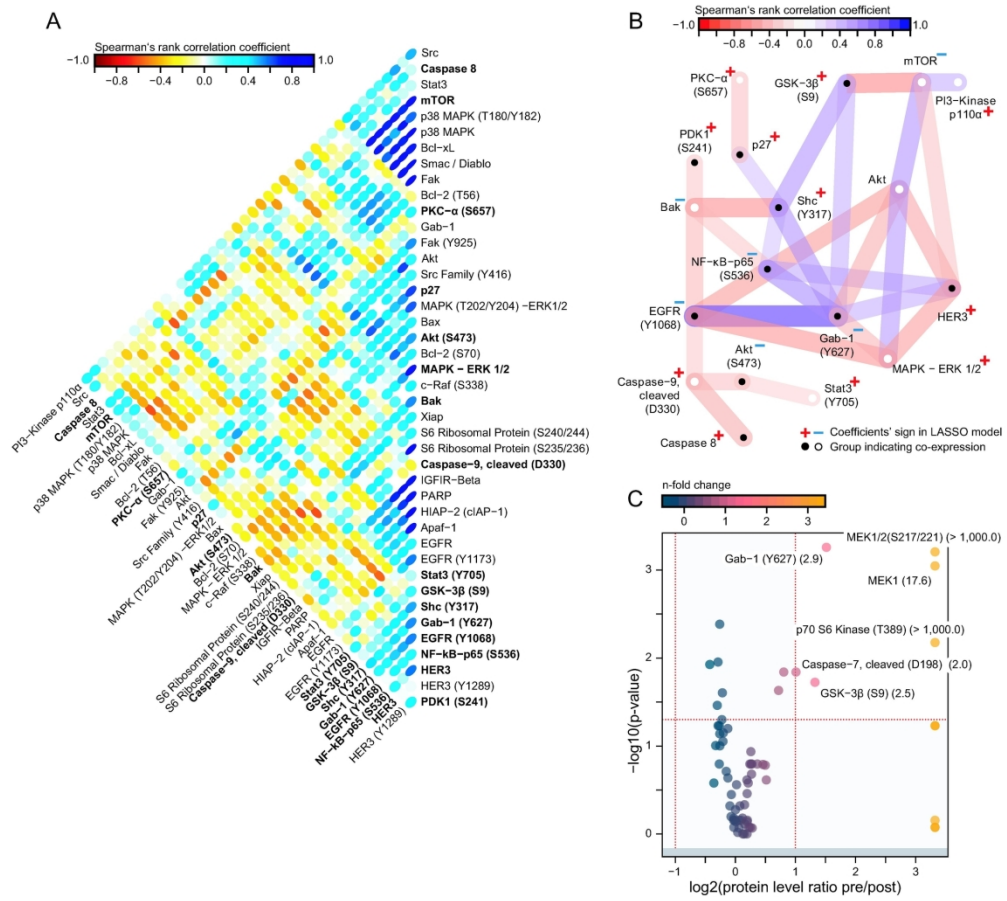
Figure 2



Protein scores indicating proteins' association to the PDX models' response to cetuximab. Proteins' scores for response to cetuximab after 3 week was calculated using PAM[27].

171x90mm (300 x 300 DPI)

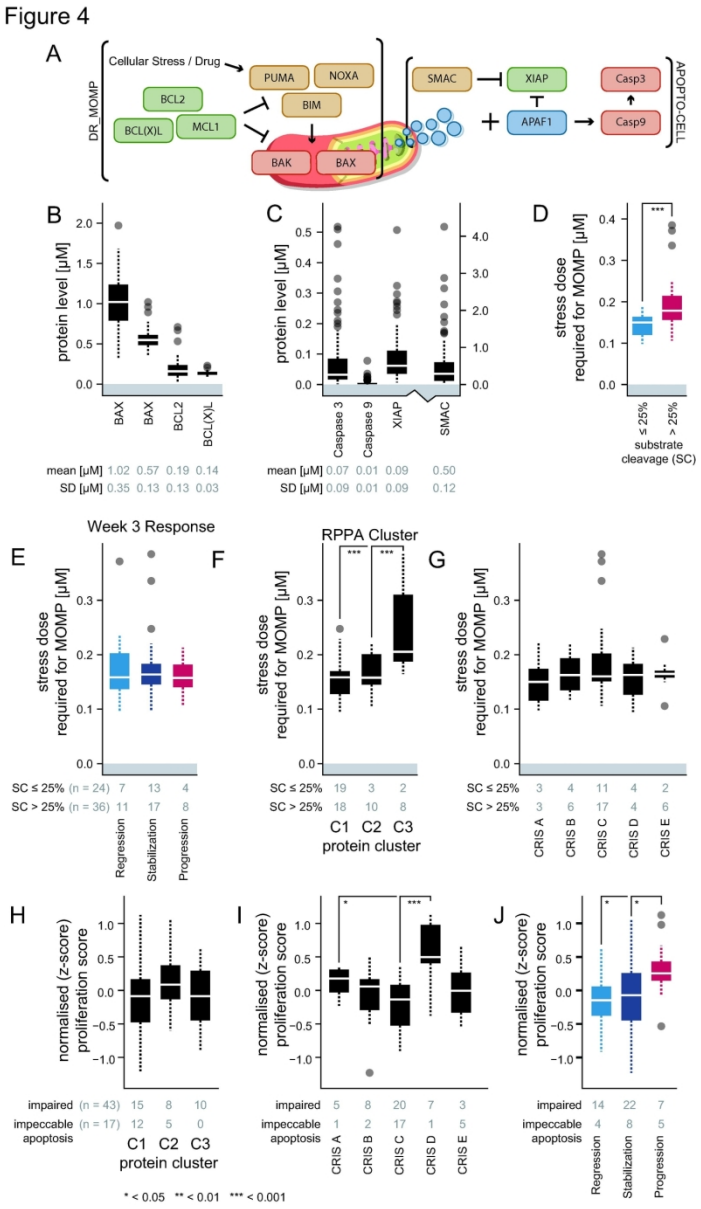
Figure 3



(A) Heatmap of Spearman's rank correlation coefficients for proteins associated with differences in response to cetuximab from Figure 2. (B) Undirected graph of proteins found to be relevant in LASSO analysis. Intensity and colour of the edges indicate the correlation coefficient of (A). Grouping based on the signs of the correlation coefficients and signs of the coefficients found by LASSO are indicated black & white nodes and plus & minus icons, respectively. (C) Protein found to be differential expressed in PDX models after treatment with cetuximab, based on pairwise comparison and Benjamin & Hochberg adjusted p-value. Dashed red lines indicate 0.05 significance threshold for p-value, and 2-fold or 1/2-fold protein level. The protein marker names and n-fold differences (treated to un-treated) in brackets were added for proteins passing all thresholds.

187x174mm (300 x 300 DPI)

Figure 4



(A) Simplified illustration of the apoptotic signalling modelled in DR_MOMP and APOPTO-CELL. Absolute protein levels normalised to HeLa cells were measured using RPPA and used as input for (B) DR_MOMP and (C) APOPTO-CELL. Calculated DR MOMP values against (D) APOPTO-CELLs' calculated substrate cleavage class with (E) differences in response to cetuximab, (F) RPPA protein cluster C3 and (G) CRIS. Calculated proliferation against (H) protein clusters, (I) CRIS and (J) response to cetuximab.

124x215mm (300 x 300 DPI)